

THIS WEEK

EDITORIALS

SUPPORT STAFF Tell your technicians you appreciate their efforts **p.528**

WORLD VIEW We must count the drug cheats to keep sport clean **p.529**

THUMBS-UP Ancient bone could have held tools much earlier **p.531**



Human history defies easy stories

The discovery of part of a 55,000-year-old human skull in Israel will help to answer some questions about our species' evolution — but it shows that the tale is complicated.

When modern humans spread out of Africa and across Eurasia 60,000–40,000 years ago, they replaced all other members of the human family, and laid the foundations for the modern world. Who were these ancestors? They left few fossils, and fewer answers.

A new piece of the puzzle is reported this week on *Nature's* website: a partial skull found in Manot Cave in northern Israel that has been dated to around 55,000 years ago (I. Hershkovitz *et al.* *Nature* <http://dx.doi.org/10.1038/nature14134> (2015) and page 541). The skull, which has a distinctive 'bun'-shaped occipital bone (the lower-back region), resembles those of modern humans found in Europe, dating to the Upper Palaeolithic starting around 50,000 years ago.

Where does the find fit in? Beware simple answers, and, indeed, simple questions. There is a temptation when discussing human evolution to reconstruct it as a narrative, in which successive species evolved to be more like us, and the more like us they became, the more likely they were to migrate to other parts of the world and replace pre-existing forms.

There are at least four things wrong with this. The first is its rather imperialist framing, in which evolution and replacement can be justified after the fact as a kind of manifest destiny.

The second is that it dismisses any extinct species as inferior and therefore of secondary importance.

The third is that it assumes the existence of an arrow of progress, in which species always evolve towards ourselves, a mistaken view that is too welcoming of spurious conceits such as 'missing links', and unwilling to countenance odd side branches such as *Homo floresiensis*, the peculiar, dwarf hominin (member of the human family) that lived in Indonesia until relatively recent times (see nature.com/hobbit10).

The fourth, and arguably the most important, is that it misrepresents the extreme fragmentation of the fossil record, something that Charles Darwin recognized, with his usual perspicience, as a 'difficulty' with his theory of evolution by natural selection. Darwin was (as usual) selling himself short. That evolution has happened is no longer in doubt: the shared chemistry and structure of all life, from the meanest microbe to the furriest feline, would be testament to that, even had no fossils ever been found.

Fossils offer more than concrete proof that evolution happened. They reveal a wealth of organic forms that no longer exist. The only species of hominin alive today, as far as we know, is our own, *Homo sapiens*. But this sole estate hides a large number of extinct forms, each of which contributed to Earth's ecology in its own particular way. If the present epoch is unusual, it is in the presence of just one species of hominin. A mere 50,000 years ago, there were at least four different species. There are very likely to have been more.

Homo sapiens first appears in the fossil record around 200,000 years ago in Ethiopia, albeit in a distinctly archaic form. The earliest fossil is not the same as the earliest member of a species — *H. sapiens* is probably much older than this. Archaic forms of our species outside

Africa first appear around 90,000 years ago, in the Levant. Another 45,000 years or so were to pass before our species made it to southeastern Europe, where it appeared amid a spectacular flourish of technology and what we would instantly recognize as art.

What happened between 90,000 and 45,000 years ago, a period ten times the length of recorded history? Only the fossils can tell us, and they are few. It seems that the earliest modern humans got to the Levant and no farther. Mount Carmel in Israel hosts caves, such as Qafzeh and Skhul, where *H. sapiens* remains appear in levels older than those occupied by Neanderthals, *Homo neanderthalensis*. The replacement of our own species by Neanderthals seems to be an affront to our prejudices. So how did humans eventually make it to Europe?

The partial *H. sapiens* skull from Manot Cave goes some way towards providing an answer, as well as hinting at how complicated our early history might have been. It looks much more modern than skulls from Qafzeh and Skhul. It is also much younger, suggesting that the hominin was closer, genetically and evolutionarily, to the earliest known European representatives of our species. This skull, the simple answer would suggest, represents modern humans poised to expand out of Africa and colonize the rest of the world.

Here comes the 'but'. Our modern genomes contain Neanderthal DNA. At some point, our ancestors bred with Neanderthals before they became extinct. Does the Manot skull represent that moment? We simply do not know. Welcome, Manot skull, to messy reality. ■

"What happened between 90,000 and 45,000 years ago? Only the fossils can tell us, and they are few."

Senate vs science

A few Republicans agreeing with basic climate research is not an environmental victory.

US President Barack Obama challenged his conservative climate critics in the annual State of the Union address to Congress on 20 January, arguing that they cannot shy away from modern climate science. A day later, pushed to take a position, 15 Republicans voted in favour of an amendment affirming the idea that humans have a role in climate change. Five went a step further, voting for a Democratic amendment stating that human activity "significantly contributes to climate change". And this is progress?

Although both amendments attracted a majority of the US Senate, neither achieved the 60-vote threshold required for approval. These votes are of course purely symbolic, but political types are already busy

reading the tea leaves for the 2016 presidential elections. For some, the fact that any Republicans, however few, felt compelled to endorse basic climate science is a positive sign that the party is once again worried about how the issue of climate change will play with US voters. We can only hope that it will at last get the attention it deserves in a major US election, but it is hard to get too excited.

The five Republicans who voted in favour of the Democratic amendment that made the strongest connection between human activity and climate change deserve credit for doing so. But the flip side is that 49 out of the 54 Republicans in the Senate voted against an amendment that merely states mainstream scientific theory, as vetted by countless researchers, studies and assessments over the course of more than a quarter of a century. And 39 refused to agree to a statement that linked human activity and climate change in any way. Moreover, it is not clear that any of the Republicans, or indeed many of the Democrats, are prepared to actually do anything significant about it.

The upshot is that little has changed. Obama has started to bypass Congress to push forward with his own climate regulations wherever possible, and he is right to do so (see page 535). If there is any criticism to be laid at his feet, it is not that he has been too ambitious with his regulatory powers, as suggested by Republicans, but that he has not been ambitious enough. His administration could certainly be more aggressive with its planned rules for power-plant emissions, as well as with methane regulations it is developing for the oil and gas sector. These regulations will help to determine whether the United States can capitalize on the shift from coal to natural gas and renewables, such as wind and solar, that has helped to reduce the nation's emissions in recent years.

For their part, Republicans have focused their energy on the Keystone XL oil pipeline from the Canadian tar sands to the US Gulf Coast, with leadership in both houses of Congress putting legislation approving it at

the top of their agenda. Environmentalists have done the same, arguing that Keystone represents a step in the wrong direction that will merely drive up greenhouse-gas emissions by promoting the development of a dirty energy source. The reality is that the pipeline, on its own, would not have a significant impact on either the US economy or the global climate.

It will be up to Obama to decide whether the pipeline is in the national interest, once the state department finishes its review of the project.

“Obama has started to bypass Congress to push forward with his own climate regulations, and he is right to do so.”

The president has said that the pipeline will benefit Canadian oil producers rather than US consumers, given that petrol prices — already lower than they have been in a long time — are driven by the international oil market. He has also said that he will approve the project only if it does not “significantly exacerbate” the problem of carbon pollution.

In the end, Obama has plenty of wiggle room in terms of how he defines both ‘national interest’ and ‘significant exacerbation’. There are surely better places to invest from a public perspective, but there are also better ways to guide private investments, including oil pipelines. One of them is to enact comprehensive climate legislation that clarifies the cost of carbon and the basic economics for all energy and infrastructure investments. That he has not done this is Obama’s biggest failure on the environmental front.

All is not lost. If the United States can continue to reduce its own emissions and help to secure meaningful action abroad, then historians may yet look back at Obama’s presidency as a turning point in the battle against global warming. One thing, however, seems clear enough: the president’s environmental legacy will not be determined by his decision on the Keystone XL pipeline. ■

Technical support

Technicians are often under appreciated, but without them there could be no research.

An old trick for book reviewers who have little material with which to judge the temperament of the author is to scrutinize the acknowledgements. Usually raw and unedited, the way these few pages of thanks are presented — gushing, self-centred or brief — can often say as much about the writer as the preceding 300 pages. The same is true for the process of science. Beneath the polished exterior of published academic papers and university press releases lies another world. And it is a world that can be glimpsed, more often than not, in the brief acknowledgements of a PhD thesis.

Alongside the praise (through gritted teeth?) for a (largely absent?) academic supervisor and the earnest gratitude showered on parents, spouses and pets for pastoral support, there is usually a list of thanks for Angela, Juan, Denise, Samuel, Ernie and a directory of other essential first-named extras. This cast of thousands is made up of the support staff and lab technicians who work behind the scenes to hold up the entire research enterprise, and who rarely get the attention they deserve.

On page 542, *Nature* makes a small effort to address this common oversight. A News Feature places a handful of these support staff front and centre, and offers details on not just their surnames, but also their crucial role. They might have more eye-catching job descriptions than many of their colleagues. But they represent an army of essential workers who are just as valuable and just as deserving of thanks.

The featured four all have very different occupations. Sarah Davis creates laboratory glassware; Jim Harrison collects venom from deadly snakes; Bill Klimm sifts the seas for squid and other inhabitants of the deep; and Dawn Johnson keeps the digital wheels turning

in a global bioinformatics archive. What they have in common is their close ties with the researchers they assist, and their remarkable and specialized skills.

Given that technical and support staff are such an important pillar of academic life, it is perhaps surprising that so little academic attention has been paid to their lot — and whether they are content with it. In 2011, researchers at King’s College London did publish a rare survey of skills and training in the United Kingdom, which raised a series of red flags (see go.nature.com/n74jsb). Technical staff are exposed on the front line when funding cuts bite: numbers working in university departments had decreased across the disciplines, both in absolute terms and relative to the number of academics and students whom they are expected to support.

One academic said: “We’re skating on thin ice — if people are away ill, or on a conference, or on training ... it’s a nightmare. If the academic department is an engine, then technicians are the engine oil that keeps the department running smoothly. Low technician numbers now mean that the department is in danger of seizing up.”

University managers should take note: the report warned that the increasing trend for centralizing services and technical support could weaken the bond between academics and technicians, and so threaten research. For example, shared mechanical workshops, formed by consolidating the facilities of several departments to save money, are unpopular and demoralizing. “University managers sometimes seem not to appreciate the vital contribution that workshop technicians make to research,” the report said. “It is important to highlight the scope for centralisation to generate problems.”

We know that PhD students appreciate the efforts of support staff, but do more senior scientists? Almost certainly. But do the technicians know that? Tell them! Do it today. Print out this editorial and pin it up in break rooms and on staff notice boards. Let technicians everywhere read the following: Angela, Juan, Denise, Samuel, Ernie — and all the rest — we salute you. ■

➔ NATURE.COM
To comment online,
click on Editorials at:
go.nature.com/xhunqv

UNIV. COLORADO



Gather data to reveal true extent of doping in sport

Drug cheats will not be tackled properly until anti-doping agencies do more to assess the scale of the problem scientifically, says Roger Pielke Jr.

How many elite athletes take performance enhancing drugs? Sporting bodies say that it is a very small minority. But a documentary broadcast in Germany last month suggested a much higher figure. Several Russian athletes claimed that nearly all of their colleagues dope, and with the knowledge of officials. The World Anti-Doping Agency (WADA) immediately launched an investigation, which is expected to report this year.

Science helps to keep sport clean by developing tests to screen athletes for banned substances. Bodies such as WADA and the US Anti-Doping Agency (USADA) say that they are doing all they can to deter doping. But they have so far neglected to carry out a simple scientific analysis of how widespread the problem is. Or if they have, they have not published the results. This makes it impossible for the rest of us to assess whether anti-doping policies are working.

Drug testing in sport, as currently implemented, might catch the occasional cheat and could deter others, but these results do little to help design an anti-doping strategy, and to independently assess whether it works. For that, we need to know whether the number of athletes doping is going up or down. And to do that, we need a reliable measure of what proportion of athletes dope. The problem — and the best way to manage it — is very different if 1% of athletes dope than if 50% of them do.

Although the stated goal of anti-doping agencies is to prevent prohibited drug use, they simply do not gather the data to enable evaluation of how effective their policies are. This is despite sporting bodies across the world spending an estimated US\$350 million on drug testing each year.

Estimating the number of elite athletes who dope is straightforward, and perfectly suited to the tools of science. Determining this number is much easier than other efforts by scientists to quantify unknowns, such as estimates of the number of planets in the Galaxy or whales in the sea. In probability-speak, it is a ball and urn problem: how do we determine how many black balls there are in an urn that contains 1,000 white and black balls if we can sample only a small number?

To assess the prevalence of sports doping, such an analysis needs two things: a reliable estimate of the total population of elite athletes and a proper randomized testing protocol. The first is readily available. For instance, at the London 2012 Summer Olympics, nearly 11,000 athletes participated from more than 200 countries. Each country conducted Olympic trials with its own pool of registered, domestic competitors seeking to qualify for the games. For the second requirement, because screening every athlete over a year is impractical, anti-doping agencies could carry out randomized tests designed to support estimates of the prevalence of doping alongside

existing testing programmes at a marginal cost.

Current doping tests are anything but random, at least in a statistical sense. Some athletes are tested several times, others not at all. In 2013, USADA says that it conducted 9,197 tests on 4,640 athletes. Decisions on which athletes were subjected to these tests were determined 'strategically', it says. The number of positive tests, then, cannot be used to say anything about a broader population. The same is true for existing global statistics. WADA says that it tested 176,502 samples (not individual athletes) in 2013, and that 1% gave 'adverse analytical findings' (AAFs).

But such a red flag does not necessarily mean that doping has occurred, because some athletes have exemptions for prohibited substances, for instance. Nor do the data allow for the matching of AAFs to sanctions against athletes. So of the 176,502 samples, what does a 1%

AAF actually mean? It is impossible to say, and that is the problem.

Why has there been no effort to quantify the problem of doping in sport (or if it has been done, why is it not published)? Evidence suggests that the leaderships of these organizations do not want to know the true extent of doping or their effectiveness in regulating it. In 2012, Richard Pound, the first president of WADA, oversaw an agency committee called Lack of Effectiveness of Testing Programs. The committee's report concluded that within the sports community, "there is no general appetite to undertake the effort and expense of a successful effort to deliver doping-free sport".

WADA, created after a major drug scandal in cycling in the late 1990s, is unique in that it is overseen by governments in partnership with non-governmental sports organizations, and operates under the provisions of a United Nations treaty. In principle, this signifies a public responsibility and expectations of accountability to stated goals. The UN treaty gives these agencies legitimacy, and thus no excuse not to be transparent.

In my opinion, anti-doping agencies suffer from a sort of institutionalized blindness that has been characterized by Steve Rayner, who studies science and civilization at the University of Oxford, UK, as the "social construction of ignorance". This is a strategy that organizations use necessarily to make their way in a complicated world. Organizations also create zones of ignorance to 'manage uncomfortable knowledge', and this can sometimes lead to dysfunction.

In the case of doping in sport, uncomfortable knowledge includes the possibility that doping among athletes is much more prevalent than is recognized and that anti-doping programmes are not very effective. But without a proper effort to gather the data, we just don't know. ■

Roger Pielke Jr is director of the Center for Science and Technology Policy Research at the University of Colorado, Boulder, Colorado, USA. e-mail: rpielkejr@gmail.com

**ANTI-DOPING
AGENCIES
SUFFER
FROM A SORT OF
INSTITUTIONALIZED
BLINDNESS.**

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/w2czqz

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

AGEING

Less cancer protein, longer life

Mice live longer — and seem to age more slowly — if they express lower levels of a cancer-promoting protein called MYC.

High levels of MYC favour tumour growth, but some expression of the protein is required for survival. John Sedivy of Brown University in Providence, Rhode Island, and his colleagues studied the effects of low MYC expression in mice. Mice with only one copy of the *Myc* gene lived 15% longer than those with two copies of the gene, although development and reproduction in the two groups were the same.

Mice with a single copy of *Myc* had a faster metabolism, and less severe age-related conditions such as osteoporosis or the thickening of the heart tissue.

Cell <http://doi.org/znb> (2015)

BIOMECHANICS

Bird's flight captured in a box

Researchers have measured the aerodynamic forces of a bird flying inside a box.

Until now, the aerodynamic lift achieved by free-flying animals has only been estimated using models. David Lentink of Stanford University in California and his team built an enclosed device to directly measure forces generated by a bird's wings during flight. With each flap, moving air exerts a force on the walls of the box, which is captured by sensors. The signals were synchronized with those from a high-speed camera, which records a bird's flight from one side of the enclosure to the other.

The researchers confirmed previous findings that each

downstroke of a bird's wings generates enough force to lift twice the animal's body weight into the air. The device could be used with other animals and free-flying robots, says the team.

J. R. Soc. Interface 12, 20141283 (2015)

CHEMISTRY

Sodium explosion caught on camera

Chemists have scrutinized a classic piece of bench chemistry — the explosion that happens when sodium metal hits water — and revised the thinking of how it works.

On contact with water,

the metal produces sodium hydroxide, hydrogen and heat, which was thought to ignite the hydrogen and cause the explosion. To delve into this, Pavel Jungwirth at the Czech Academy of Sciences in Prague and his team used high-speed cameras to capture the reaction of a drop of a liquid alloy of sodium and potassium with water at room temperature.

They found that spikes of the metal shoot out from the droplet just 0.4 milliseconds after it enters the water — too fast to have been expelled by heat. Computer simulations revealed that sodium atoms at the surface of a small cluster each lose an electron within picoseconds. The positively

would affect) river volume worldwide between 1930 and 2030. The team found significant changes to existing water flow in rivers such as the Parana River in South America (pictured). The biggest future effects would arise from dams being planned for the Amazon basin.

The models could help engineers to reduce the environmental effects of new dams.

Environ. Res. Lett. 10, 015001 (2015)



JORGE SAEZ/AP

HYDROLOGY

Dams reshape the world's rivers

Dams have altered 48% of all river flow worldwide. And if all dams planned for the next few decades are built, that proportion will nearly double.

Günther Grill of McGill University in Montreal, Canada, and his team developed two ways to analyse how dams break up and regulate river flow. They calculated how 6,374 existing dams and 3,377 proposed ones affected (or

charged ions rapidly repel each other, causing the explosion, while the protruding metal spikes generate new surface area that drives the reaction.

Nature Chem. <http://dx.doi.org/10.1038/nchem.2161> (2015)

CLIMATE SCIENCE

Big swings in weather to come

Weather extremes could become more common as the climate warms this century, because extreme cooling events in the Pacific Ocean are predicted to occur more often.

La Niña events occur when the equatorial Pacific cools, causing droughts and floods

worldwide. Wenju Cai of the Commonwealth Scientific and Industrial Research Organization in Aspendale, Australia, and his colleagues analysed the occurrence of significant La Niñas and related El Niño events from 1900 to 2009, simulated under rising concentrations of greenhouse gases in the atmosphere. The researchers found that the number of extreme La Niña events increased from one every 23 years to one every 13 years in the twenty-first century.

Most of the severe La Niñas will follow severe El Niños, resulting in wide, annual swings between opposite extreme weather events, the authors suggest.

Nature Clim. Change <http://doi.org/zph> (2015)

PALAEOANTHROPOLOGY

Ancient hands built for tools

The hands of hominins that lived about 3 million years ago were capable of clutching tools.

The first tool-using hominin is widely believed to have been *Homo habilis* — known as the handyman — in part because its appearance in the fossil record 2.4 million years ago coincides with the earliest stone tools. To search for earlier signs of tool use, a team led by Matthew Skinner and Tracey Kivell at the University of Kent, UK, analysed the composition of the hand bones of *Australopithecus africanus* fossils from South Africa, which are between 2 million and 3 million years old. The ends of *A. africanus* metacarpal hand bones (**pictured**), which form the palm, resembled those of later toolmakers such as *Homo sapiens* and Neanderthals.

The team concludes that *A. africanus* could forcefully grip objects using an opposable thumb. **Science** 347, 395–399 (2015)



ECOLOGY

Pumas feel the fear near humans

Female pumas that live near human populations hunt more often but spend less time eating their prey than do those in less populated areas.

Humans can cause declines in wildlife populations, but their effect on animal behaviour is less well understood. Justine Smith and her colleagues at the University of California, Santa Cruz, tagged 30 pumas (*Puma concolor*) in California and tracked their movements in areas with four different densities of human housing. They found that at kill sites near the most densely populated areas, female pumas spent 42% less time consuming their prey than those in the least populated regions. To compensate, the females in the more developed habitats killed 36% more deer.

Fear of humans is probably driving this behavioural change, which could have further ecosystem effects, such as boosting scavenger populations and even compromising the reproductive health of female pumas, the authors speculate. **Proc. R. Soc. B** 282, 20142711 (2015)

ENVIRONMENTAL SCIENCE

Methane escapes from major city

The ageing pipeline infrastructure of Boston, Massachusetts, is leaking natural gas — mostly methane, a potent greenhouse gas — at more than double the rate of previous estimates.

Atmospheric methane levels had plateaued but have been growing worldwide since 2007, for reasons that are unclear. Kathryn McKain at Harvard University in Cambridge, Massachusetts, and her colleagues monitored

SOCIAL SELECTION

Popular articles on social media

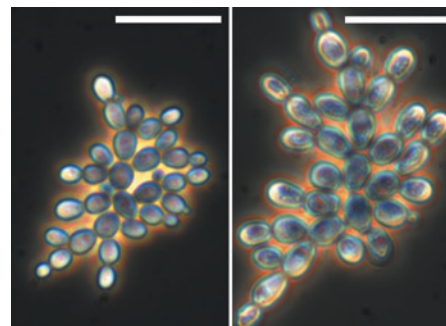
Celebrating beauty in science writing

Not many people read research articles for the snappy writing. But Stephen Heard, an ecologist at the University of New Brunswick in Fredericton, Canada, argues in a blogpost (go.nature.com/a2xh1m) that scientific writing could be more readable and even elegant, an observation that set off a widespread social-media reaction. Heard wrote that researchers should try livening up their scientific prose to attract and keep more readers. Isabelle Côté, a marine ecologist at Simon Fraser University in Burnaby, Canada, tweeted: “Let’s put some whimsy, humour and beauty in scientific writing.” Anthony Caravaggi, a conservation biologist at Queen’s University Belfast, UK, tweeted: “I’d love to see less turgidity & more charm.”

➔ **NATURE.COM**
For more on popular papers:
go.nature.com/uwgzik

methane levels at four locations in Boston for a year. They also used a model of atmospheric processes to determine methane emissions. They found that 60–100% of the emitted methane was from the city’s natural-gas system, and that the Boston region is losing about 2.7% of its natural gas: 2–3 times more than other estimates.

Cities that consume natural gas could be a bigger source of atmospheric methane than was previously thought. **Proc. Natl Acad. Sci. USA** <http://doi.org/zpk> (2015)



By mathematically modelling the way that clusters break off, the authors conclude that this way of growing makes the cells in each cluster genetically similar. This allows natural selection to act on the clusters rather than on individual cells, speeding up multicellular evolution. A mutation in a gene encoding the protein ACE2 causes the clusters to form.

After 60 days of selection (400 generations), the yeast evolved bigger cells (**pictured**, right; scale bars are 50 μm) compared with those at 14 days (left). The results show how a single mutation can create multicellular clusters and set the stage for the future evolution of organismal complexity. **Nature Commun.** 6, 6102 (2015)

➔ **NATURE.COM**
For the latest research published by Nature visit:
www.nature.com/latestresearch

EVOLUTION

How yeast go multicellular

A genetic mutation in single-celled yeast turns it into a multicellular organism — hinting at how multicellularity might have evolved.

William Ratcliff at the Georgia Institute of Technology in Atlanta and his co-workers studied a strain of yeast (*Saccharomyces cerevisiae*) in which the daughter cells remain attached to the mother cells after dividing, resulting in multicellular ‘snowflake’ yeast.

SEVEN DAYS

The news in brief

PEOPLE

Breeders convicted

An Italian court found three employees of a lab-animal breeding company guilty of animal cruelty on 23 January, sentencing them to prison terms of between 12 and 18 months. The court said that the three had mistreated beagle dogs and had put some animals down without good reason. Marshall BioResources, the US company that operates the Green Hill facility near Brescia, Italy, had previously been cleared of wrongdoing, but prosecutors pressed charges against Green Hill director Roberto Bravi and three other workers, one of whom was later found innocent. Even if the verdict is upheld by higher courts, the sentences could still be converted to fines or probation. See go.nature.com/sbcdvz for more.

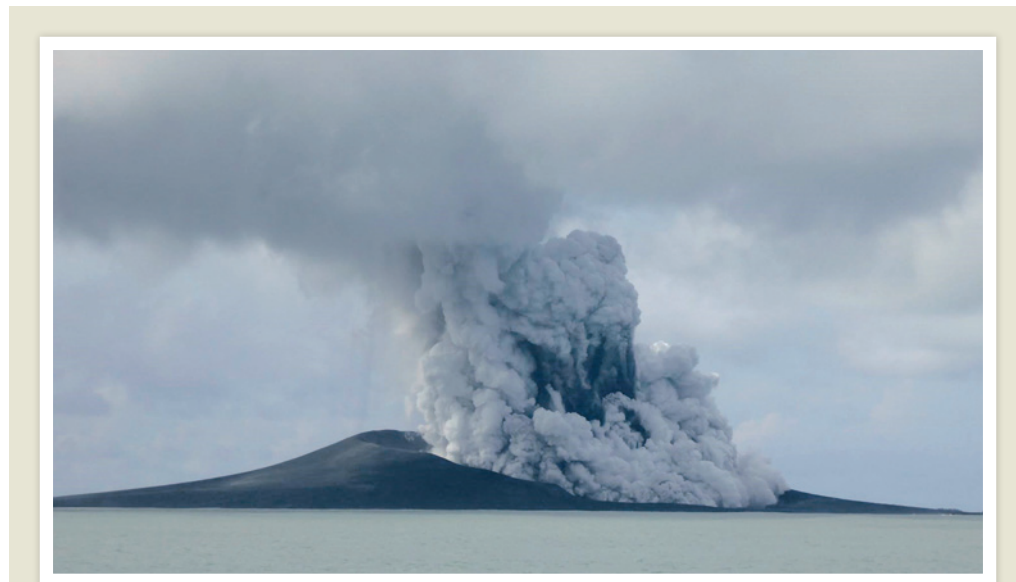
Israel frees scientist

A Palestinian physicist who had been detained by the Israeli military was released on 22 January. Imad Ahmad Barghouthi of Al-Quds University in Jerusalem was on his way to a scientific conference on 6 December when he was arrested while attempting to cross from the West Bank to Jordan. Israel's reasons for administrative detention are normally kept secret, but Barghouthi claims that he had been jailed because his profile picture on Facebook showed him wearing a green scarf, the colour of the Hamas movement. See go.nature.com/qbgliy for more.

POLICY

US medicine effort

US President Barack Obama announced a Precision Medicine Initiative during the annual State of the Union



NZ MINISTRY OF FOREIGN AFFAIRS & TRADE/AP

New island born from volcanic eruption

A volcanic eruption in the Pacific archipelago of Tonga has created a new island. Steam and ash began rising last month between the islands of Hunga Tonga and Hunga Ha'apai, about 60 kilometres north of the kingdom's main island of Tongatapu. The newborn isle

breached the surface around 20 December. In mid-January volcanologists measured it at about 1.8 kilometres long and 100 metres high. A similar island formed during the volcano's last eruption in 2009, but was washed away by ocean waves within weeks.

address on 20 January. "I want the country that eliminated polio and mapped the human genome to lead a new era of medicine — one that delivers the right treatment at the right time," he said. The effort will aim to match genomic and other data with patient health records to discover new treatments and tailor existing ones. Obama is expected to request hundreds of millions of dollars to fund research for the initiative at the US National Institutes of Health, which will have a key role. See page 540 for more.

Vaccine price jump

The medical charity Médecins Sans Frontières in Geneva, Switzerland, called on drug companies to cut the price of vaccines in a 20 January

report. A full programme of childhood vaccines now costs 68 times more than it did in 2001, on the basis of the best prices available to low-income countries. Much of the price jump is due to expensive vaccines that protect against bacterial pneumococcal diseases, which have been rolled out in many low- and middle-income countries in recent years.

Sea-science report

The US National Science Foundation (NSF) should immediately slash spending on marine hardware, says a report released by the National Research Council on 23 January. The report, which is intended to guide US oceanography for the next decade, says that basic ocean research at the NSF

is losing out to the rising costs of infrastructure. It calls for a 20% cut to the operating budget of the Ocean Observatories Initiative, a marine monitoring effort due for completion in May. The initiative is currently expected to cost up to US\$59 million per year. See page 538 for more.

India energy deal

US President Barack Obama and Indian Prime Minister Narendra Modi pledged on 25 January to expand cooperation on clean energy and tackling climate change. During Obama's visit to New Delhi, the two leaders agreed to work towards an international climate agreement as well as a separate plan to reduce emissions of

hydrofluorocarbons, a class of powerful greenhouse gases that are often used as refrigerants. India has committed to installing 100 gigawatts of solar energy by 2022, but has made no concrete promises about curbing greenhouse-gas emissions.

Marijuana research

A leading medical academy has called for the US government to loosen restrictions on marijuana, so that researchers can study its potential medical benefits for sick children. In a 26 January policy statement, the American Academy of Pediatrics said that marijuana should no longer be classed alongside heroin and ecstasy as a schedule 1 drug with no accepted medical use. It argues that the drug should be relabelled as schedule 2, placing marijuana in the same category as potentially addictive painkillers such as morphine and codeine. Several federal agencies are already considering such a change.

RESEARCH

Tiger boom

Tigers are mounting a comeback in India. Last week, the country's government announced a 30% increase in the population of Bengal



tigers (*Panthera tigris tigris*; pictured), from 1,706 individuals counted in a 2010 census using thousands of camera traps, to 2,226 recorded last year. A 2004 survey documented just 1,411 tigers. The Indian government attributed the increase to efforts to limit poaching and minimize human encroachment on the feline's habitat. An estimated 70% of the world's tigers now reside in India.

Fish from the cold

A marine ecosystem has been discovered beneath the Ross Ice Shelf, which extends hundreds of kilometres off the coastline of Antarctica. After drilling through 740 metres of ice, researchers sent a remotely operated vehicle to explore the area around the borehole on 16 January. The vehicle photographed fish and marine crustaceans known as amphipods thriving in the pitch dark, -2°C waters — the nearest to the South Pole that

such an ecosystem has been found. See go.nature.com/jomocy for more.

EVENTS

No climate hoax

The US Senate resoundingly affirmed the existence of global warming during a 21 January debate over legislation that would authorize construction of the Keystone XL oil pipeline. Lawmakers voted 98–1 to adopt a 16-word, non-binding amendment: “It is the sense of the Senate that climate change is real and not a hoax.” Another amendment, to pin some of the blame for climate change on human activities, failed by 50–49. See go.nature.com/tqjdn for more.

Doomsday clock

The *Bulletin of the Atomic Scientists* pushed the hands of its Doomsday Clock two minutes closer to midnight on 22 January — symbolizing that the planet is closer to

global disaster. The physicist-founded magazine, based in Chicago, Illinois, cited unchecked climate change and the threat of outsized nuclear-weapons arsenals as the reasons for the shift. The clock now stands at three minutes to midnight: the closest it has been since 1984, but further away than in 1953 after the hydrogen bomb was first tested, when the clock stood at two minutes to midnight.

PUBLISHING

Open journals

The University of California Press plans to establish two open-access journals, it announced on 20 January. A ‘mega journal’ called *Collabra* will publish research articles in the biomedical, life, environmental and social sciences — and it will be unusual in paying reviewers and editors for their time. They will have the option of donating the money to fund other papers published by *Collabra* or giving it to their own institution to pay for open access. The second journal, *Luminos*, will publish research monographs.

BUSINESS

Space investments

SpaceX added US\$1 billion to its coffers with financing from Google and the investment firm Fidelity, it announced on 20 January. The company, of Hawthorne, California, is one of several aerospace firms currently carrying cargo to the International Space Station for NASA, and it is working towards flying US astronauts there in the future. Company founder Elon Musk wants to send astronauts to Mars one day. For now, however, the investments will support SpaceX's plans to develop reusable rockets and manufacture fleets of small satellites, which are likely to provide Internet access.

➔ NATURE.COM

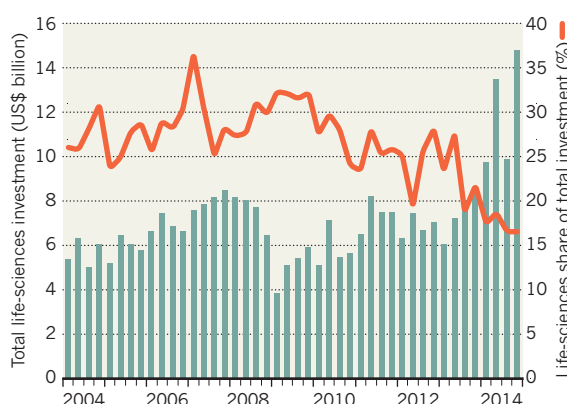
For daily news updates see:
www.nature.com/news

TREND WATCH

Venture capitalists invested more in the life sciences last year than they have since 2008, according to data released on 16 January by PriceWaterhouseCoopers and the US National Venture Capital Association. Investment in biotechnology and medical-devices companies rose by 29% on 2013, driven by deals in fields such as cancer immunotherapy. Public investors also embraced health care, with 102 companies launching an initial public offering in 2014, compared with 54 in 2013.

BOOM TIME FOR BIOTECH

Venture capitalists, who have showered money on software start-ups in recent years, gave life sciences a shot in the arm in 2014.



► regulations and the growth of renewable energy sources, especially wind power.

But first, the regulations must withstand inevitable industry lawsuits intended to weaken or overturn them. “The power-plant rule is really pivotal to Obama’s legacy, and it is going to face tough legal scrutiny,” says David Victor, director of the Laboratory on International Law and Regulation at the University of California, San Diego. If it succeeds, he adds, “it could reverse the position of the United States internationally”.

Although Obama was unable to secure climate legislation during his first term, his administration did achieve gains in the wake of the recession. It secured billions of dollars in stimulus funding for clean energy, efficiency measures and green infrastructure, as well as establishing significant new standards for vehicle emissions and fuel economy. That, combined with the economic slowdown and the shift away from coal in the electricity sector, means that US greenhouse-gas emissions have already decreased by around 10% since 2005.

In theory, the administration still has both the time and the means to reduce emissions

enough for the United States to meet its international commitments, says Kennedy. “They will have to take very serious action, but the tools that they have available to them should allow them to do it.”

Republicans have vowed to challenge Obama at every turn. They started the current session with a debate on legislation to approve the controversial Keystone XL pipeline, which would carry crude oil from the tar sands of Alberta, Canada, to refineries on the US Gulf Coast. The House of Representatives quickly passed a bill to approve the pipeline, but partisan disagreements have delayed a Senate vote. Obama has promised to veto the legislation.

Although the pipeline would have a small effect on global greenhouse-gas emissions, it has become a symbolic issue for both sides of the climate debate. On 21 January, Senate Democrats used the Keystone fight to confront Republicans on their views about climate by

putting to a vote declarations about human involvement in global warming (see *Nature* <http://doi.org/zpx>; 2015). Fifteen Republicans supported an amendment to the Keystone bill stating that climate change is affected by human activity, and five voted for an amendment stating that climate change is “significantly” affected by humans.

Although neither amendment passed, those votes are a sign that Republicans are feeling pressure and may warm to certain climate solutions in future, says Bob Inglis, a Republican former member of the House who heads the Energy and Enterprise Initiative, a think tank that advocates for conservative environmental solutions at George Mason University in Fairfax, Virginia. Although Inglis understands why Obama has chosen to sidestep Congress and address climate change with regulations, he says that the president still has a potential opportunity to secure his environmental legacy by striking a grand legislative bargain with his Republican opposition.

“Obama is in a box,” says Inglis, “but he could get out of that box if he were a little bit bolder.” ■ **SEE EDITORIAL P. 527**

Source: H. Sierks et al. *SCIENCE* <http://doi.org/zp2> (2015)

SOLAR SYSTEM

Philae hunt hangs in the balance

Rosetta mission would have to sacrifice other science to search for comet lander.

BY ELIZABETH GIBNEY

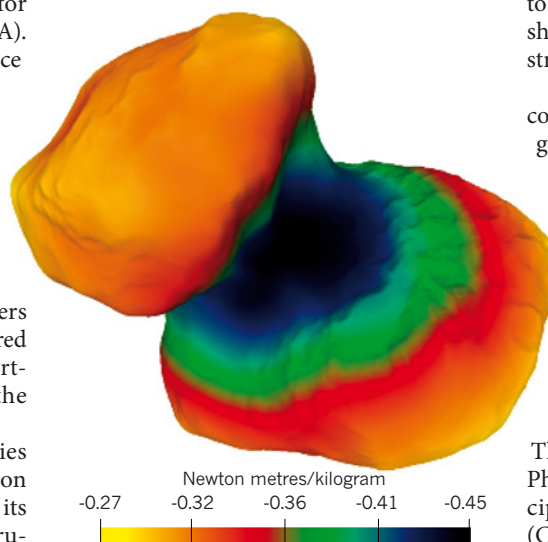
The lost space probe Philae, which made history after it landed on a comet last November, is posing a dilemma for scientists at the European Space Agency (ESA). They have what is probably their last chance to change the path of Philae’s parent craft, Rosetta, to hunt for the lander, which went missing shortly after it touched down on comet 67P/Churyumov–Gerasimenko. But the shift would also mean sacrificing some of Rosetta’s long-planned science observations.

The agonizing choice comes as the mission team published its first batch of papers from observations made after Rosetta entered into orbit around 67P last August — reporting a varied landscape and hinting at the comet’s origins.

Philae has been silent since its batteries ran out just days after its bumpy landing on 12 November. On the basis of images of its initial bounces and data from radio instruments, Philae’s position has been narrowed down to a 20-metre by 200-metre strip. But efforts to find the 1-metre-wide lander in high-resolution pictures taken by Rosetta from a distance of about 20 kilometres have so far failed.

NEARLY WEIGHTLESS

The resulting effect of gravitational potential and centrifugal forces, mapped on comet 67P/Churyumov–Gerasimenko, is revealed to be greatest on the lobes and weaker in the neck region.



Project scientists are debating whether to send Rosetta, which is still orbiting the comet, down to an altitude of 6 kilometres, over the patch where Philae is thought to be. It would

be the closest that the craft has ever been to 67P. But Rosetta has limited fuel. Any attempt to look for Philae would mean scrapping a different flyby, which would offer the chance to image the comet in a shadow-free shot that should reveal fine details about the surface structure and composition.

As the comet approaches the Sun, growing cometary surface activity in the form of jets of gas and dust also makes it increasingly risky for Rosetta to approach. If Rosetta is to stick to the original flyby plan, scheduled for 14 February, the craft will not come as close to 67P until 2016, says the mission’s flight director Andrea Accomazzo, after the comet has swung around the Sun and headed back out to space.

SURFACE SEARCH

There are scientific benefits to pinpointing Philae’s location, says Wlodek Kofman, principal investigator on Rosetta’s CONSERT (Comet Nucleus Sound-ing Experiment by Radio-wave Transmission) experiment, which is designed to send radio waves between the

► **NATURE.COM**

For the best images from Rosetta’s data haul, see: go.nature.com/rrihsj

parent craft and Philae to study the comet's interior. Not knowing the lander's exact location makes it much harder to process the data that scientists have already received and to generate accurate results, he says. Spotting the lander would also help to determine its exact location and angle, and to predict how likely it is to come back to life in the coming months as the comet nears the Sun and its solar panels begin to receive more light, Kofman says.

The decision is not an easy one, says Holger Sierks, who is principal investigator on Rosetta's OSIRIS (Optical, Spectroscopic, and Infrared Remote Imaging System) instrument.

The mission has already produced a haul of results, which were published in a series of papers in *Science* on 22 January (see *Nature* <http://doi.org/zip; 2015>). Using data from OSIRIS and the Radio Science Investigation instrument, Sierks and his collaborators calculated the gravity on the rubber-duck-shaped comet and created a map (see 'Nearly weightless') that also takes into account the centrifugal force caused by the comet's rotation (H. Sierks *et al.* *Science* <http://doi.org/zip2; 2015>). The resulting force is greatest on top of the lobes, but it is about six times weaker in the neck region, where dust can lift off more easily. The team also used the data to calculate the comet's density, finding that the body is relatively fluffy and porous — with a density of around half that of water, giving clues to its structure and strength.

The researchers described three-metre-wide pebble-like features that are found all over the comet, which they nicknamed "goosebumps". Sierks says that the shapes could hint at the size of the grains of dust and ice that first clumped together in the early Solar System before forming larger bodies. "The hypothesis is these might be the building blocks of comets," he says.

In another of the papers, OSIRIS data enabled Sierks and his collaborators to classify the geography of the comet's surface on the basis of terrain types. These include fractures, possible impact craters and an array of dunes and ripples that may have been formed by gas travelling around the surface, like wind shaping sand in a desert (N. Thomas *et al.* *Science* <http://doi.org/zip3; 2015>).

The final word on whether to send Rosetta to look for Philae rests with ESA. Kofman says that an informal vote among Rosetta scientists came down narrowly on the side of doing it. As *Nature* went to press, the agency was thought to be leaning towards sticking to its original agenda, because looking for Philae would mean too much upheaval for the mission.

If ESA decides against a mission shift to hunt for Philae, the team could still get lucky: it may find clues as to the lander's whereabouts either in existing images or in new shots taken from flybys between 20 km and 50 km away in the coming months. Finding Philae in these kind of flybys is not impossible, says Accomazzo, "but it would be sheer luck". ■

PSYCHOLOGY

Clash over 'smart unconscious'

Report examining decisions made while distracted adds to controversy about the power of the unconscious.

BY ALISON ABBOTT

If you have to make a complex decision, will you do a better job if you absorb yourself in, say, a crossword puzzle instead of ruminating about your options? The idea that unconscious thought is sometimes more powerful than conscious thought is attractive, and echoes ideas popularized by books such as writer Malcolm Gladwell's best-selling *Blink*.

But within the scientific community, 'unconscious-thought advantage' (UTA) has been controversial. Now Dutch psychologists have carried out the most rigorous study yet of UTA — and find no evidence for it.

Their conclusion, published this week in *Judgement and Decision Making*, is based on a large experiment that they designed to provide the best chance of capturing the effect should it exist, along with a sophisticated statistical analysis of previously published data¹.

The report adds to broader concerns about the quality of psychology studies and to an ongoing controversy about the extent to which unconscious thought in general can influence behaviour. "The bigger debate is about how clever our unconscious is," says cognitive psychologist David Shanks of University College London. "This carefully constructed paper makes a great contribution." Shanks published a review last year that questioned research claiming that various unconscious influences, including UTA, affect decision making².

A typical study probing UTA asks subjects to make a complex decision, such as choosing a car or a computer, after either mulling over a list of the object's attributes or viewing the list quickly and then engaging in a distracting activity such as a word puzzle. However, such studies have drawn different conclusions, with about half of those published so far reporting a UTA effect and the other half finding none.

Proponents of the theory claim that the effect is exquisitely sensitive to experimental variations, and often attribute the negative results to the fact that many research groups varied elements of the set-up, such as the choice of puzzle used for the distraction³.

Critics say that the positive results came from having too few participants in the experiments.

Psychologists Mark Nieuwenstein and Hedderik van Rijn at the University of Groningen in the Netherlands set out with their colleagues to determine which explanation was correct.

They asked 399 participants — around ten times more than the typical (median) sample sizes in other studies — to choose between either 4 cars or 4 apartments on the basis of 12 desirable or undesirable features. They incorporated the full list of conditions that UTA proponents had reported as yielding the strongest effect, such as the exact type of puzzle used as a distraction. They found that

"How we make decisions, and how we might make them better, has practical and intellectual importance."

the distracted group was no more likely than the deliberating group to choose the most desirable item.

The scientists then reanalysed 60 of the 81 experiments described in the 32 UTA papers published before April 2014. For this 'meta-analysis', they excluded experiments that had insufficient data for analysis or that deviated from conditions that are reported as likely to elicit UTA (only one of these experiments had claimed a UTA effect). They also included the results of their own study. When they applied a rigorous statistical meta-analysis, they found no significant UTA effect.

"Psychologists have historically prided themselves on their command of statistics," says psychologist Jonathan Baron at the University of Pennsylvania in Philadelphia, the editor of *Judgement and Decision Making*. But this study shows that many in the past were poorly designed. He adds: "If UTA is out there, it can't be captured in experiments designed in the lab."

Psychologist Ap Dijksterhuis at Radboud University in Nijmegen, the Netherlands, who first described⁴ unconscious-thought theory, which predicts UTA, in 2004, says: "It is certainly true that psychology has ►

improved quite a bit in recent years when it comes to analysing data. And yes, in the past, suboptimal analyses have been applied.” But he does not accept the findings of the meta-analysis. He says that it would have produced different conclusions if the researchers had included all previous UTA experiments, rather than excluding some and relying on a subset. He adds that “the evidence for UTA is growing quickly” and is widely accepted.

UTA is not the only ‘smart-unconscious’ claim to come under scrutiny. For example, experiments carried out under the “Many Labs” Replication Project, which coordinates labs internationally to repeat psychological studies in order to validate their claims, as well as several separate studies,

have challenged another psychological concept, social priming. Under social priming, certain behaviours are claimed to be modified unconsciously by previous exposure to stimuli, such as an American flag, or thinking about money⁵.

Other doubts raised about unconscious thought include its role in some types of decision making under uncertainty.

In spite of the most recent findings, Brian Nosek, a psychologist at the University of Virginia in Charlottesville who co-launched Many Labs, says that he remains optimistic about the theory underlying UTA. “I would be surprised if unconscious-thought theory did not hold up, because it fits with contemporary theories,” he says.

Shanks agrees that the debate over unconscious-thought theory is probably not over. “How we make decisions, and how we might make them better, has practical and intellectual importance,” he says. “If there is any evidence that distraction or unconscious rumination helped, we’d want to know about it — but the conclusions are so far very premature.” ■

1. Nieuwenstein, M. R. *et al. Judge. Decis. Making* **10**, 1–17 (2015).
2. Newell, B. R. & Shanks, D. R. *Behav. Brain Sci.* **37**, 1–19 (2014).
3. Strick, M. *et al. Social Cogn.* **29**, 738–762 (2011).
4. Dijksterhuis, A. J. *Pers. Social Psychol.* **87**, 586–598 (2004).
5. Klein, R. A. *et al. Social Psychol.* **45**, 142–152 (2013).

OCEANOGRAPHY

US ocean sciences told to plot fresh course

Major report calls for cuts in infrastructure funding to increase spending on science.

BY ALEXANDRA WITZE

Faced with the rising costs of going to sea, the ocean-sciences division of the US National Science Foundation (NSF) should slash what it spends on marine hardware to fund more research, says a major report by the US National Research Council. It proposes making the biggest cut to the showcase US\$386-million Ocean Observatories Initiative (OOI), which after years of construction is just months away from being finished.

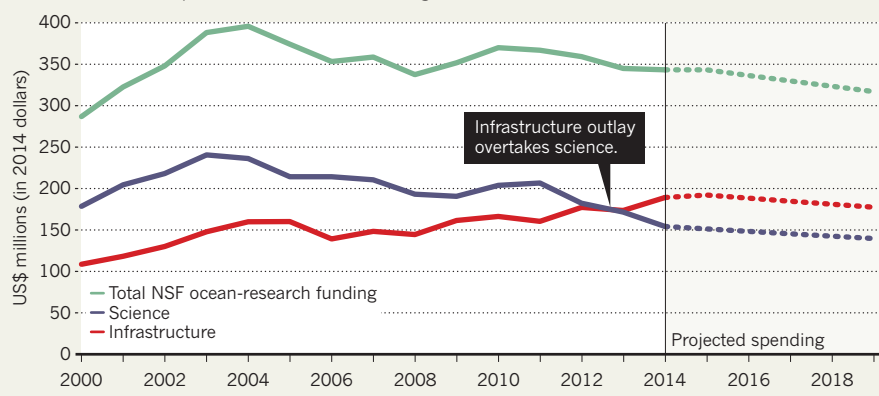
The report’s authors suggest that the NSF should cut 20% of the OOI’s operations budget, and reduce its contributions to the international scientific ocean-drilling programme and the US academic research fleet. If the agency takes that advice, it could free up enough money for US oceanographers to begin to reclaim much of their lost science, as well as expand partnerships with international researchers.

“It’s an exciting time to be in ocean science,” says Shirley Pomponi, an oceanographer at Florida Atlantic University in Fort Pierce and co-chair of the report committee. “But we need to take steps to make that better.”

US oceanography has been in trouble for a while. The US Navy paid for the bulk of the country’s academic oceanographic work until the 1960s, after which the NSF began shouldering more of the burden. But even as filmmaker James Cameron, flush with private money, explored the Pacific Ocean’s Mariana

SINKING SCIENCE

As the US National Science Foundation (NSF) has increased its spending on ocean hardware, such as ships and instruments, its funding for ocean science has fallen.



SOURCE: NRC

Trench with a handful of scientists in 2012, most research oceanographers found themselves with fewer ways to get to sea.

EIGHT PRIORITIES

Over the past decade, the NSF’s ocean-infrastructure expenses have risen by 18% — even as the ocean-science division’s inflation-adjusted budget dropped by more than 10%, to just under \$350 million annually. In 2013, the division started to spend more on infrastructure than it did on science (see ‘Sinking science’). That is when the NSF asked for outside advice on how to cope.

The report, which was published on 23 January, lays out eight science priorities for the next decade, including studies of sea-level change, marine biodiversity, earthquakes and tsunamis, and life beneath the sea floor. Unusually, it also suggests how to pay for the studies — an immediate 10% cut in infrastructure, spread unequally among three programmes, followed by a similar or larger cut over the next five to ten years. “This document gives them the flexibility to make some really hard decisions,” says Samantha Joye, an oceanographer at the University of Georgia in Athens.

The smallest suggested cut, just 5%, applies

US ocean sciences told to steer a new course

Major report calls for cuts to infrastructure, including fledgling Ocean Observatories Initiative, to increase spending on science.

Alexandra Witze

23 January 2015



Lamont-Doherty Earth Observatory

The US National Science Foundation should cut spending on research ships by 5%, a new report says.

Faced with rising costs of going to sea, the ocean-sciences division of the US National Science Foundation (NSF) should immediately slash what it spends on marine hardware, says a new report. It suggests making the biggest cut to the flagship US\$386-million Ocean Observatories Initiative (OOI), which after years of construction is just months away from being finished.

The report, released on 23 January by the US National Research Council, is likely to guide US oceanography for years to come. It is the first formal attempt to address what many researchers have grumbled about for years — that basic ocean science at the NSF is losing out to the rising costs of infrastructure.

Top picks
from **nature** news

- Science pours in from Rosetta comet

Back to basics

To get science funding back to its historical level, the report's authors suggest slashing 20% of the OOI's operations budget, and making smaller cuts to the NSF contributions to the scientific ocean-drilling programme and the US academic research fleet. If the agency takes that advice, it could free up enough money for US oceanographers to begin to reclaim much of their lost science, as well as expand partnerships with international researchers.

"It's an exciting time to be in ocean science," says Shirley Pomponi, an oceanographer at Florida Atlantic University in Fort Pierce and co-chair of the report committee. "But we need to take steps to make that better."

US oceanography has been in trouble for a while. The US Navy paid for the bulk of the country's academic oceanographic work until the 1960s, after which the NSF began shouldering more of the burden. Even as filmmaker James Cameron, flush with private money, explored the Pacific Ocean's Mariana Trench with a handful of scientists in 2012, most research oceanographers found themselves with fewer ways to get to sea.

Over the past decade, the NSF's ocean infrastructure expenses rose 18% as the ocean-science division's inflation-adjusted budget dropped by more than 10%, to just under \$350 million annually. In 2013, the division started spending more on infrastructure than it did on science. That is when the NSF asked for outside advice on how to cope.

Eight priorities

The report lays out eight science priorities for the next decade, including sea-level change, marine biodiversity, earthquakes and tsunamis, and life beneath the sea floor. Unusually, it also suggests how to pay for the studies — a 10% cut in infrastructure immediately, spread unequally among three programmes, followed by a similar or larger cut over the next five to 10 years. "This document gives them the flexibility to make some really hard decisions," says Samantha Joye, an oceanographer at the University of Georgia in Athens.

The smallest suggested cut, just 5%, applies to the NSF contribution to the 20-vessel research fleet, because ships enable more of the future science priorities. Even so, the report recommends that the agency build no more than two new 'regional class' vessels in the coming years; it had been considering building three.

The middle cut, of 10%, is recommended for the scientific ocean-drilling programme — in the guise of the *JOIDES Resolution* drillship, which has already endured a number of cutbacks.

Finally, Pomponi and her colleagues suggest that 20% should be cut

mission

- GM microbes created that can't escape the lab
- Crunch time for pet theory on dark matter

Related stories

- US Arctic research ship ready to cast off
- Marine science: Oceanography's billion-dollar baby
- Drilling hit by budget woes

More related stories

from the OOI's operations budget, which will run between \$55 million and \$59 million when it begins full operations this year. They note that the OOI is made of many components, some of which do a better job than others at studying the key science priorities.

For instance, a cabled sea-floor observatory off the coast of Oregon addresses the risks of underwater earthquakes and tsunamis. Two moveable arrays of instrumented moorings — one off the US east coast and one off the west coast — tackle questions such as regional sea-level change. But four deep-water sites — two at high northern latitudes and two in the far south — are much less crucial for the eight priorities, says the report. It argues that at least one of the southern sites could be sacrificed without much scientific loss, while others may need only to be instrumented for two or three years to understand ocean change rather than the 25-year planned lifetime of the OOI.

Seeking balance

"We didn't say the OOI was not important," says report co-author Melbourne Briscoe, an oceanographer who has previously worked for the government and now runs an environmental consulting company in Alexandria, Virginia. "It's just that it wasn't as relevant to that set of eight questions as other pieces of infrastructure were."

Officials of the Consortium for Ocean Leadership, the Washington, DC-based organization that is building the OOI for the NSF, declined to comment, saying that they needed more time to digest the report and its implications.

Richard Murray, who became director of NSF's ocean sciences division this month, says that his top priority is to figure out how to move forward with the new community advice. "We're looking at this report very, very seriously," he says. "We hear very clearly the committee's recommendation that something needs to be done."

Mitchell Lyle, an oceanographer with Oregon State University in Corvallis who was not involved in the report, says that the field needs to stop quibbling over how to spend ever-smaller slices of the budgetary pie, and instead take a fresh look at what US oceanography can really afford. "Funding levels are now dropping below a sustainable level to support the number of ocean scientists that we now have," he says. Oceanographers might be better served by looking at more dramatic options for saving federal research dollars, he says, such as possibly closing some of the current oceanographic institutions.

Nature **517**, 538–539 (29 January 2015) doi:10.1038/nature.2015.16780

Related stories and links

From nature.com

- **US Arctic research ship ready to cast off**
27 May 2014
- **Marine science: Oceanography's billion-dollar baby**

25 September 2013

- **Drilling hit by budget woes**

25 September 2013

From elsewhere

- **Decadal Survey of Ocean Sciences 2015**

For the best commenting experience, please login or register as a user and agree to our Community Guidelines. You will be re-directed back to this page where you will see comments updating in real-time and have the ability to recommend comments to other users.

1 comment

[Subscribe to comments](#)

Mitch Lyle • 2015-01-25 11:58 PM

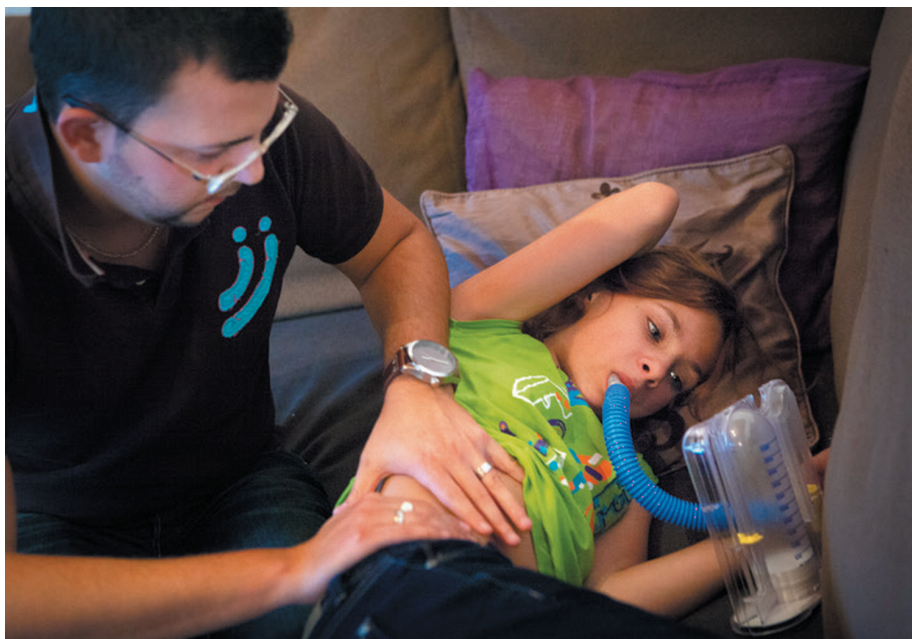
As the person quoted in the last paragraph, I would like to clarify my conversation—the point I was trying to make is that oceanographic institutions are already being damaged by the current and projected funding climate. The most likely path that the report describes is continued flat funding for ocean sciences. While oceanography can survive on flat funding for a while and the report describes ways it can, ocean science has been weakening over the last decade and will continue to weaken as the field expands but funding does not. Despite important new science to pursue, new technology will not reduce the expense of data collection to the point that we can afford new people for doing the science. OOI, IODP, and previous long term plans for OCE arose from the assumption at the turn of the century that there was strong bipartisan support for doubling the NSF budget. The NSF budget did not double and the budget in ocean sciences shrank by 20%. At the present time we are also losing skills and expertise through retirement and loss of positions, and are not hiring new scientists and technicians rapidly enough to maintain a healthy science. This trend will continue unless additional funding appears. NSF cannot close oceanographic institutions, and I didn't suggest that it should. I did point out that one of the current weaknesses in the field is that the numbers of institutions that exist are too great for the dollars available, and all institutions will tend to grow weaker and expertise will tend to become more diffuse under flat funding. There is a strong tendency now for technical groups to lose critical mass of people and lose functionality. If we seriously believe that there will be no new funding, we need to rethink how we do oceanography. How do we get healthy again?

See other News & Comment articles from *Nature*

Nature ISSN 0028-0836 EISSN 1476-4687

© 2015 Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.

partner of AGORA, HINARI, OARE, INASP, CrossRef and COUNTER



Some children with cystic fibrosis are eligible for targeted treatments.

HEALTH CARE

Precision-medicine plan raises hopes

US initiative highlights growing focus on targeted therapies.

BY SARA REARDON

With the pipeline of conventional drugs drying up, researchers are increasingly attempting to customize treatments on the basis of a person's genetics or environment. Now the US government wants to get in on the act.

During his State of the Union address to Congress on 20 January, President Barack Obama announced a programme called the Precision Medicine Initiative. "I want the country that eliminated polio and mapped the human genome to lead a new era of medicine — one that delivers the right treatment at the right time," he said.

The White House is remaining tight-lipped about the details of the programme, declining to answer questions from *Nature* — as is the US National Institutes of Health (NIH), a key partner in the effort. But Kay Holcombe, senior vice-president for science policy at the Biotechnology Industry Organization (BIO) in Washington DC, says that her conversations with the NIH suggest that the initiative will seek to match genome information with many other data types, such as health records and blood-test results.

The agency seems to have been planning

the effort for some time, listing precision medicine as one of its four priorities in its 2015 budget proposal; another was big data. Other government agencies are also expected to participate, as may some private companies. Further details, including

"My personal attitude is always to try to collaborate, rather than duplicate and compete in an inefficient way."

the cost, are likely to trickle out as Obama prepares his budget request for fiscal year 2016, which is due to be released on 2 February. A major question is whether the plan will run alongside or merge with a similar proposal being discussed by members of the US House of Representatives' Energy & Commerce Committee. The committee's 21st Century Cures plan seeks to speed up the translation of research advances into treatments, and personalized medicine is one potential element of the effort. Law-makers are expected to release a first draft of that proposal shortly.

Both the White House effort and the House plan would be extremely expensive, but they might not be as difficult to carry out as they

first seem. Rather than recruiting all of their participants anew, Holcombe says that both initiatives could collect data and recruit participants from ongoing longitudinal studies. These include the Million Veteran Program at the US Department of Veterans Affairs, which seeks to understand how genes affect health, and the NIH's 67-year-old Framingham Heart Study at Boston University in Massachusetts, which aims to identify risk factors for heart disease.

SCANT DETAILS

If the federal programme takes the form of a public-private partnership, then private insurance companies and health systems could contribute data as well. David Ledbetter, chief scientific officer at Geisinger Health System in Danville, Pennsylvania, says that his company might be willing to join such an effort. Geisinger, a network of hospitals and clinics, aims to recruit up to 200,000 of its 3 million customers to have their exomes — parts of the genome that code for proteins — sequenced and integrated with their health records. The company now has completed sequences from about 20,000 people, and it is preparing to provide each person with an analysis of his or her health risks.

"My personal attitude is always to try to collaborate, rather than duplicate and compete in an inefficient way," Ledbetter says.

Still, standardizing data collection and patient recruitment across the country will be extremely difficult, especially if ongoing studies are rolled into the effort. Such complexities sank the NIH's 100,000-person National Children's Study, which sought to track environmental influences on children's health; the agency cancelled the project last month after 14 years of delays.

Informed consent and data security will present additional challenges. The roll-out of the UK National Health Service's care.data project, which would make health information from most patients in England available for research, has been delayed for several months for this reason.

Nevertheless, with personalized medicine in vogue, studies are likely to continue to grow in both number and magnitude, and in both the public and private sectors. The Precision Medicine Initiative could once again pit NIH director Francis Collins, who headed the Human Genome Project, against his old private-sector rival, Craig Venter. Last March, Venter launched a company called Human Longevity in San Diego, California, with the goal of sequencing one million human genomes by 2020. The effort is gaining steam: on 14 January, Venter announced that his company would be sequencing tens of thousands of genomes for Genentech, a biotechnology company based in South San Francisco, California, that is searching for new drug targets. ■

AMELIE-BENOIST/BSIP/CORBIS



ISRAEL HERSHKOVITZ, OFER WARDER & OMRY BARZILAI

Manot Cave, near Israel's Sea of Galilee, was rediscovered when a bulldozer opened an entrance to it in 2008.

ARCHAEOLOGY

Neanderthals gain human neighbour

Cranium discovery shows that Homo sapiens was living in Middle East 55,000 years ago.

BY EWEN CALLAWAY

A 55,000-year-old incomplete skull found in Israel may belong to a human group that interbred with Neanderthals. Discovered deep in a cave by amateur speleologists, the partial cranium also fills a major gap in the fossil record of *Homo sapiens*' journey from Africa to Europe.

"Here we actually hold a skull of a human being that was living next to the Neanderthals," says Israel Hershkovitz, the leader of a study published today in *Nature* (I. Hershkovitz *et al.* *Nature* <http://dx.doi.org/10.1038/nature14134>; 2015). "Potentially he is the one that could interbreed with the Neanderthals," says Hershkovitz, who is a physical anthropologist at Tel Aviv University in Israel.

Genome studies of Neanderthals (*Homo neanderthalensis*) and of both ancient and contemporary *H. sapiens* suggest that the two species interbred somewhere in the Middle East between 50,000 and 60,000 years ago (Q. Fu *et al.* *Nature* **514**, 445–449; 2014). But the problem with this idea is that no remains of anatomically modern humans have been discovered in the Middle East from this crucial period, after *H. sapiens* left Africa and before it colonized Europe and Asia.

In 2008, a bulldozer clearing land for a development near the Sea of Galilee in northern Israel revealed an opening to a limestone cave that had been sealed for more than 15,000 years. Amateur speleologists were the first to explore the cave, and they spotted the battered bone — the top portion of a skull — resting on a ledge. The Israel Antiquities Authority soon



The 'skullcap' fills a gap in the path that modern humans took as they spread out of Africa.

launched a complete survey of Manot Cave, finding buried stone tools at several spots that are still being excavated.

The skull was unquestionably from *H. sapiens*, says Hershkovitz: it was similar in shape to those of earlier African and later European humans. A patina of calcite coated the fragment, and the researchers used radioactive uranium in the mineral to date the bone to about 55,000 years old. That means that "the Manot people are probably the forefathers of the early Palaeolithic populations of Europe", Hershkovitz says.

The Manot people are also a leading candidate for the humans that bred with Neanderthals — exploits that have given all of today's non-African humans a sliver of Neanderthal heritage. The Manot Cave is not far from two other sites that held Neanderthal remains of a similar age. "The southern Levant is the only place where anatomically modern humans and Neanderthals were living side by side for thousands and thousands of years," Hershkovitz says. The ultimate proof would be to look for the presence of Neanderthal ancestry in DNA from the skull, but the region's

balmy temperatures mean that ancient DNA is unlikely to have been preserved.

Jean-Jacques Hublin, a palaeoanthropologist at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, agrees that the chances of recovering DNA from the skull fragment are slim. But he hopes that further excavations will find human remains that have stayed cool enough to still contain DNA. These digs might also connect the skull to stone tools and other relics of daily life, which could strengthen the Manot skull's link to early Europeans. The artefacts uncovered so far are thought to be much younger than the skull. "We have a skull, and we have a site where there is some archaeology, but there is no link between the skull and the archaeology. It's a bit annoying," Hublin says.

"This specimen is really important and exciting, as — assuming the dating is correct — it shows for the first time that modern humans existed in the Near East at the same time as Neanderthals," says Katerina Harvati, a palaeoanthropologist at the University of Tübingen in Germany. "Until now we had no evidence that the two even coexisted in this region during this time period. So this is a crucial piece of the puzzle." ■ [SEE EDITORIAL P.527](#)

CORRECTION

The News Feature 'Laser focus' (*Nature* **517**, 430–432; 2015) gave the wrong amount for the funding behind the Center for Adaptive Optics. The grant was in fact for around US\$40 million over 10 years.

FROM I. HERSHKOVITZ ET AL. *NATURE* [HTTP://DX.DOI.ORG/10.1038/NATURE14134](http://dx.doi.org/10.1038/nature14134) (2015)

NOT YOUR AVERAGE TECHNICIAN

Research relies on unsung heroes working behind the scenes — and some of them have rather unusual jobs.



The glass-blower in the bush

BY MICHAEL HOPKIN

The West Australian town of Jarrahdale (population 1,082) seems an unlikely place to go if you need to get your hands on some highly technical glassware in a hurry. Turn off the main street with its tavern, general store and logging museum, and the road quickly becomes dirt punctuated by sun-faded letter boxes, wonky fences and dusty driveways.

But it is down one of these driveways that you'll find Sarah Davis, who has been running a scientific glass-blowing business since 2010. Working from her garage, she provides local researchers — mostly university chemists in nearby Perth — with handmade flasks, tubes, condensers and bespoke items that don't even have a name.

"If they want a simple condenser, I can whip that up in half an hour," says Davis, referring to the glass tube used to cool hot vapours. "I get people ringing up saying, 'I've broken this,' and generally I get it out for them the next day." For scientists who live in one of the world's remotest cities, this makes Davis an extremely useful person to have around. The alternative is to wait at least six weeks for orders to be made and shipped from Sydney. "Sometimes she comes in after a couple of days and says, 'I've finished,' and I say, 'Already?'" says Grant Cope, who orders from Davis as part of his job as stores officer for the chemistry department at Curtin University in Perth.

Many big research institutions have their own scientific glass-blowers — and that is what Davis was doing until five years ago, working as the in-house glass-blower on the University of Western Australia



(UWA) campus in Perth. But in 2010, when she was laid off in a round of university cutbacks, she decided to go it alone, putting her outbuildings into service as her workshops.

"What could be more Australian than working in your garage and seeing a kangaroo come hopping down the drive?" she says. This is, in fact, routine. She also shares the garage with two possums that like to take naps in the rafters when the temperature creeps past 40 °C, as it tends to do in summer. But in the worst of the weather, Davis is less likely to be toiling over hot glass: she is a volunteer firefighter and is regularly called up to deal with bush fires.

"If they want a simple condenser, I can whip that up in half an hour."

The rustic setting belies the fact that Davis's craft is a highly technical practice, bearing very little resemblance to traditional glass-blowing. For a start, there is not much blowing involved.

She works with borosilicate glass, which unlike standard glass, can withstand temperatures of 300 °C, as well as corrosive chemicals

and high pressures. She heats and softens the glass over a gas flame, then uses a variety of tools to work it into shape. Perhaps most important is the glass-blowing lathe, with two spindles facing one another, both turning at precisely the same speed. On a day in December, with blowtorch in one hand and safety glasses firmly on (hot borosilicate glass gives off a dangerously intense orange glare, not to mention lots of ultraviolet radiation), she carefully attaches a section of glass to the end of a long tube mounted on the lathe, rounding it off to create a test tube the size of her arm. She uses a similar process to make her flasks and other more specialized glassware.

To finish off, Davis bakes her wares at 560 °C in an annealing oven, smoothing out stress points that could otherwise break the glass. With



MATT DEVLIN

The snake milker

BY KELLY RAE CHI

In nearly four decades collecting deadly snake venom, Jim Harrison says, he has been bitten “only eight times”. And although he remembers each one vividly, tallying them up on his fingers can be tricky. An Indian cobra (*Naja naja*) mangled his right little finger 12 years ago, leaving it curled and increasingly sore until he had surgery to repair it. A bite from a desert horned viper (*Cerastes cerastes*) dissolved part of the bone in his left middle finger. Two other fingers, although functional, bear the scars of his profession.

All this is par for the course when you nurture lethal snakes for science. Harrison and his wife, Kristen Wiley, run the Kentucky Reptile Zoo (KRZ) in Slade, which Harrison opened in 1990 as a research and education centre. It houses 1,600 snakes from more than 100 species, and it is one of just a handful of places around the world producing snake venom for biomedical research.

Snake venoms contain a complex cocktail of enzymes and other substances that help to immobilize or digest prey, and which are of great interest to scientists. Drugs used to treat hypertension have been modelled on substances in venom that drastically lower the blood pressure of prey, for example. Other proteins in ▶

diamond saws, tube cutters, lathe and oven, Davis estimates that her set-up is probably worth around half a million Australian dollars (US\$400,000), and as a result she prefers to keep a low profile — even in a quiet town. “I don’t tend to have clients come and see me, I don’t have a web page; it’s all word of mouth and previous clients.”

She has plenty of those, garnered over a 20-year career that started when, as a newly qualified lab technician in Perth, she landed a job that included a glass-blowing traineeship. Davis admits that she had never heard of scientific glass-blowing before that. George Koutsantonis, a chemist at the UWA, describes her components as “vital” for his research on pyrophoric chemicals, which ignite spontaneously if exposed to air. “It’s not the sort of thing you can buy off the shelf,” he says. Davis’s strangest commission so far has been from some intrepid zoologists who asked her to make a glass funnel to hold over a dolphin’s blowhole in the hope of catching a sample for analysis. “I never got to see it in action,” she says.

These kinds of weird and wonderful commissions are a lot rarer now. Thanks to financial pressures, only a handful of Australian universities still have an on-campus glass-blower — researchers have to order off-the-shelf glassware, and are less likely to request customized parts if they have to pay freelance glass-blowers out of tight budgets.

Even counting those still plying their trade off-campus, there are only 25 scientific glass-blowers left in Australia and New Zealand, says Davis. “There are just two of us in Western Australia that do it — the other guy is getting to retirement age. Hopefully I’ve got another 25 years left in me, but the chance of training someone is probably not there. It’s a dying art.” ■

➔ **NATURE.COM**
For videos of the
glass-blower and
squid collector, see:
go.nature.com/k5oule

DAVID STEPHENSON/CATERS NEWS



► venom have been used to identify and study specific signalling molecules in the nervous system. And venoms are needed to develop antivenoms. The KRZ sells about 1,400 grams of venom per year.

Wiley and Harrison “provide a tremendous service, because most of us don’t have time to be zookeepers”, says Steven Aird at the Okinawa Institute of Science and Technology Graduate University in Japan, who has studied venom. “They really become not just suppliers, but almost collaborators in a sense.”

Harrison’s fascination with snakes and other reptiles took hold when he caught a garter snake at the age of six. Throughout childhood, he read voraciously on reptiles and amphibians; at 16, he worked on an alligator farm.

Harrison started keeping venomous snakes as a hobby. He learned about venoms and extraction from books, including those written by Sherman Minton, a prominent herpetologist in Indianapolis with whom Harrison eventually became friends. Minton connected Harrison with others interested in venoms, and soon Harrison began to milk king cobras (*Ophiophagus hannah*) for university researchers.

Harrison never believed that he could have a career involving snakes, so he became a police officer instead. But he continued extracting venom in a home laboratory equipped with a centrifuge to purify venom and a lyophilizer for freeze-drying it. At 26, after getting mown down by a stolen car while trying to make an arrest, Harrison’s heart stopped. He decided that policing was too dangerous, so he retired early and dedicated his career to snakes. Since then, snake bites have stopped Harrison’s heart three more times.

DEADLY DISPLAY

These days, Harrison and Wiley divide the work of running the reptile zoo. Wiley, who did an internship at the KRZ in 1998, manages the zoo’s educational programmes, reads the scientific literature and attends conferences to stay current on venoms and work out whether to breed a particular species that year.

The actual milking falls to Harrison, who for liability reasons is the only staff member at the KRZ who does it. In front of a group of goggle-eyed schoolchildren, he demonstrates his technique on a monocled cobra (*Naja kaouthia*), a species that put him in hospital on life support after a bite in 2012. He pulls the 1.2-metre-long, dishwater-grey specimen onto a padded mat and pins its head down with the flat part of a long metal hook.

Harrison grabs the cobra behind its head. As it reveals its fangs — a natural response to threat — Harrison plants them through a sheet of plastic film stretched across a funnel. He uses his thumb and a partially missing forefinger to massage the muscle supporting its venom glands. He will do this on between 600 and 1,000 snakes per week. If everything goes as it should, he says, then milking snakes is methodical — “boring”, even. In fact, according to data that a physician friend gathered on him, Harrison’s heart beats faster when he is driving to the supermarket than when he is milking.

Stephen Mackessy at the University of Northern Colorado in Greeley says that the KRZ’s reputation and knowledge of venoms sets it apart. Some companies provide repackaged venoms, but the provenance of these products, which can matter greatly in research, is uncertain at best, he says. Wiley says that much of this comes down to understanding the animals, which she and Harrison breed themselves, but also obtain from zoos and universities. “We attempt, as much as we can, to provide the locale and the origin information to the researcher,” says Wiley.

Harrison says that the benefits for medical researchers — and for society — make him willing to take his daily calculated risks. “I don’t plan on slowing down,” he says. “I will keep extracting until I die.” ■



The squid collector

BY ELIE DOLGIN

On a blustery morning in late October, the wind is blowing up quite a swell — enough to make this reporter heave his breakfast into the briny deep — but Bill Klimm is unperturbed. The 78-year-old fisherman sits calmly in his captain’s seat, arms folded, staring straight ahead at the choppy waters off the coast of Martha’s Vineyard, Massachusetts, as his boat, the *Gemma*, travels southwest.

Klimm and his co-captain, Dan Sullivan, are heading to Menemsha Bight in search of longfin inshore squid (*Doryteuthis pealeii*). These squid are prized for their giant nerve fibres, which allow biologists to study neurotransmission in exquisite detail. For the past 18 years, Klimm has been collecting these and other saltwater specimens for scientists at the Marine Biological Laboratory (MBL) in Woods Hole, Massachusetts, and elsewhere around the world.

From invertebrates such as sponges, worms, sea stars, urchins and anemones to several fish species and some plants, the creatures have a wide range of habits and dwelling places, but Klimm knows where to find them. And if he does not, he has a network of local fishermen that he can tap for advice.

David Remsen, who manages the Marine Resources Department at the MBL, tells Klimm what to catch on the basis of the orders he receives from scientists. He says that a good specimen collector needs intuition for the local seas and the skills to maintain the boats that navigate them. Klimm has it all. “He knows the waters, he knows the equipment, and he takes ownership of both,” says Remsen.

Klimm’s knowledge of marine biology runs deep, too. “If you want to understand something about the squid life cycle, you will learn more in ten minutes talking to Bill than you will spending a week talking to so-called experts,” says Joseph DeGiorgis, a squid neurobiologist at Providence College in Rhode Island, and an adjunct faculty member at the MBL.

A LIFE AT SEA

H. William Klimm III was born with Cape Cod fishing in his blood. His grandfather was a fisherman and lobsterman who owned a boatyard in Hyannis Harbor, Massachusetts. His father was a commercial fisherman operating out of Falmouth, who collected squid for the MBL as a sideline for 45 years — until the age of 88.

DANIEL COJANU PHOTOGRAPHY





Klimm himself started fishing commercially when he was 23 years old. He caught cod, flounder, swordfish and lobsters for 30 years, until a boat fire cast him ashore in 1990. He fixed boats in Boston for five years before landing the MBL position. At the age of 60, after decades of long trips at sea, Klimm was finally home every night after work. “My wife calls it a toy job,” he says.

David Bodznick, who studies the neurobiology of behaviour at Wesleyan University in Middletown, Connecticut, has spent summers at the MBL for more than 30 years, researching electrosensing in skate. “You could really tell when [Klimm] came on board that changes had been made,” he says. For example, Klimm installed new reels and altered the nets to minimize damage to the squid and other animals. “The whole operation became more efficient,” says Bodznick.

When Klimm and Sullivan reach their

destination at Menemsha Bight, they drop a large net into the water, tow it along for 25 minutes, pull in the line and sort through the catch — all without exchanging a word. “We do it so many times that we don’t have to talk about it,” Klimm remarks afterwards. Large squid (those

“You will learn more in ten minutes talking to Bill than you will spending a week talking to so-called experts.”

25 centimetres long or more) go in one bucket; medium in another. Any small or damaged animals get tossed to the squawking seagulls overhead.

A majority of the hundred or so squid collected today will be used to train neurosurgeons attending a week-long teaching course at the MBL. Some will go to the nearby Woods Hole Oceanographic Institution, where scientists are investigating the effects

of ocean acidification on squid physiology, and 10–20 go to a visiting researcher at the MBL, Yuyu Song, who is studying how misfolded proteins affect neurotransmission in the squid’s giant synapse. “The *Gemma*, her captain and the MBL collecting expeditions are all very dear to me since a big part of my research would have been impossible without them,” says Song, a neuroscientist normally based at the Yale School of Medicine in New Haven, Connecticut.

Back in port, Klimm talks about what he does with his free time, gesturing across the dock to his “play boat”, the *Sea Dog IV*. That is where he and his wife can be found most weekends in the summer, tooling around Martha’s Vineyard and the Nantucket Sound. “For years and years and years I’ve done that — stepped from one boat to the other,” he says. “It’s kind of stupid, I suppose, but that’s what I do. That’s what I do.” ■



TOBY SMITH/GETTY IMAGES REPORTAGE



The data mechanic

BY EWEN CALLAWAY

When Dawn Johnson opens the doors into her workspace, the first thing you notice is the roar. The noise comes from whirring fans, which are required to cool the towering stacks of 16 computer servers that form walls of black and silver. Bundles of multicoloured cables, as thick as small trees, trail upwards like an electrical rainbow.

"If anything goes wrong, I'll be the first port of call," Johnson says, standing beside a toolbox the size of a shopping trolley. "I'll rip 'em to bits and find out."

Computational biologists the world over rely on Johnson to do that, even though most do not know her. That's because Johnson is a computer-hardware engineer at the European Bioinformatics Institute (EBI) in Hinxton, UK. The servers that she keeps running hold one of the world's most extensive collections of molecular databases — from an archive of DNA-sequencing data to the leading repository of protein structures. The machines that she and her colleagues maintain hold a whopping 60,000 terabytes of data, and people at around half a million unique Internet addresses use these data each month. A blip in availability is not an option. "It's imperative that it's there 24/7," says Johnson.

For Johnson, bearing the weight of the bioinformatics world on her shoulders was particularly burdensome late last year. Besides the centre in Hinxton, the EBI data had been spread across another two locations in London, but a contractor change meant that they had to move to a single location in a nearby town — and Johnson had to coordinate it. She and a small team of fellow engineers had to ensure that there was adequate space, power and cabling for the move, which involved roughly

9,500 computers connected by 850 power cables and 3,400 network cables. "The complexity of that is hair-raising," she says, with a relaxed shrug. Still, the move went "incredibly well," says Steven Newhouse, head of technical services at the EBI — with Johnson playing a crucial part in coordinating the logistics. The success meant, of course, that the researchers who rely on the EBI never so much as noticed. "Very few scientists appreciate the size of the computing infrastructure that they depend on nowadays," says Newhouse.

When she is not at her desk dealing with the ins and outs of such projects, Johnson spends her time in the Hinxton data centre, which the EBI shares with the neighbouring Wellcome Trust Sanger Institute. Johnson and her colleagues install, maintain and repair the machines that feed the centres' seemingly insatiable hunger for data storage — which is projected to reach 2 exabytes (2×10^{18} bytes, or 2 million terabytes) by 2016. There are occasional emergencies. Several years ago, a cooling-system failure forced Johnson to rush into work on a Saturday to keep the servers from overheating. She spent a stressful weekend getting the centre back online, so as to minimize the disruption to researchers.

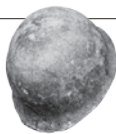
Computers were not the first machines that Johnson learned to rip apart. "My father's a mechanic and an engineer, and so I was always in the garage with him fixing and tinkering with cars, and that was really what I wanted to do," Johnson says. "But it was 1979 when I left school. They just didn't hire lady mechanics." She went into secretarial work at a firm in Cambridge, UK, that sold and serviced computers for businesses. After a few years, her boss asked her what she wanted to try next, and she opted to work as a computer engineer. She was the only woman on the team.

"When I was in the field, I was a novelty, I guess. But it was quite good. All the guys wanted to help me get on and succeed, and all the women saw me as sort of a stand for women's lib and rights and stuff and were really on my side as well," she says. Even now, "I don't meet many other women in my career, which is a shame".

Johnson's move into the bioinformatics world happened by chance. In the 1990s, she was doing contract work on mainframe computers at the Sanger Institute, which had a leading role in the Human Genome Project. She remembers a celebration to mark the completion of a draft human-genome sequence. "I saw that happening and thought I would like to be a part of that," says Johnson. A hardware-engineer job opened up five years ago, and she jumped at the opportunity. "It's great when I drive into work and hear people on the radio talking about the latest studies," she says. "I'm very proud and lucky to be part of it." ■ [SEE EDITORIAL P.527](#)

COMMENT

SOCIOBIOLOGY Two takes on altruism, its roots and ramifications **p.550**



ENERGY Fracking forecasters respond on shale-gas estimates **p.553**

SUSTAINABILITY Year of soil draws much-needed attention to sustainability crisis **p.553**

OBITUARY Donald Metcalf, blood-proliferation pioneer, remembered **p.554**

AIMEN EL SAHLI/REUTERS/CORBIS



Rock art thought to be about 4,000 years old in Libya's Tadrart Acacus mountains was vandalized in 2009.

Save Libyan archaeology

Until violence eases and fieldwork can resume, fund research in labs, museums and on computers, urges **Savino di Lernia**.

Libya is a hotspot for research into the human past. The Sahara, the largest hot desert in the world, was once green and hosted until a few thousand years ago the biggest freshwater lake on Earth¹. Some depictions of crocodiles and cattle engraved and painted on the walls of rock shelters in the Sahara date back 9,000 years.

The desert is also a laboratory for investigating links between past climate changes and developments in human history^{2,3}. These include the dispersal of modern humans across Africa about 130,000 years ago⁴, the oldest evidence⁵ of milking in Africa around 5200 BC and the establishment

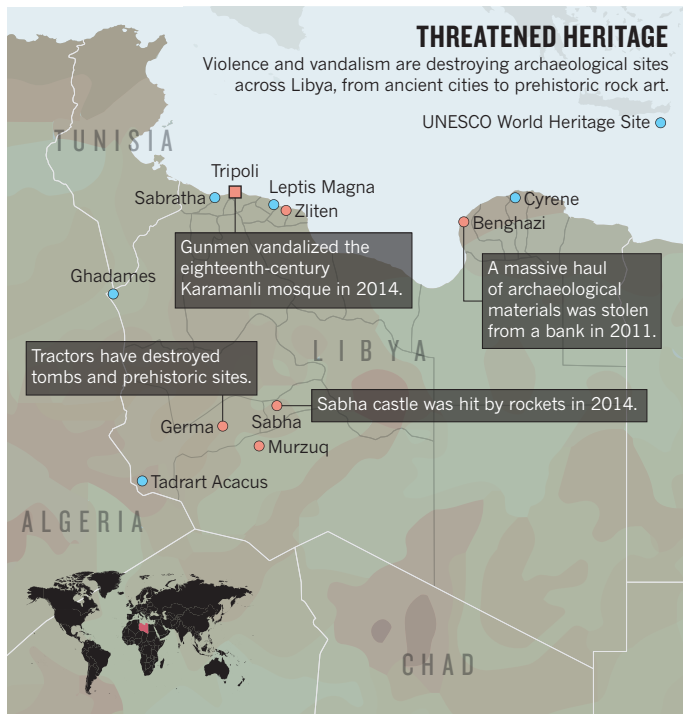
of the first Saharan state⁶ during the first millennium BC.

Archaeological fieldwork in Libya is at a standstill. Four years after the Arab Spring and the February 2011 Libyan revolution that ended the regime of Muammar Gaddafi, violence remains rife. Recent escalations in fighting have injured and killed people and damaged the nation's cultural heritage, infrastructure and free press. Libyan monuments have been seriously damaged, including the Karamanli mosque, built in 1738 in the capital, Tripoli, and Islamic tombs that date to between the tenth and twelfth centuries at Zuwila, near the west-central town

of Murzuq. This, along with concerns about the illicit trafficking of cultural materials, led Irina Bokova, the director-general of the United Nations Educational, Scientific and Cultural Organization (UNESCO), to call for greater protection of Libyan cultural heritage in November last year.

The destruction of archaeological sites in Syria, Iraq and Afghanistan — to name but a few other war-torn countries — are part of the same picture: what does not comply with militant revolutionaries' aims is expendable or must be destroyed.

I have worked in Libya since 1990. My last field trip to the Messak plateau in the ►



The ancient Greek city of Cyrene is at risk because of nearby construction.

► southwest ended abruptly in February 2011 with an emergency evacuation on a military aircraft. Before the revolution, I spent three months each year in the desert studying the prehistory of the Messak and nearby Tadrart Acacus mountains, which lie close to the border with Algeria, famous for their 9,000-year-old rock art. Since then, scientific and cultural relations between Libya and the international community have stagnated. Archaeological tourism — a major source of revenue and jobs for locals such as the Tuareg and Tebu people, the two major Saharan ethnic groups in Libya — has stopped.

Even though a discussion of cultural heritage might seem out of place in a country devastated by civil war, I argue that scientific research in the region must not be abandoned. As UNESCO recognizes, culture has a powerful role in “building social cohesion and contributing to reconciliation and peace”. We must continue to nurture skills, trust and knowledge about our shared past. Until fieldwork in the region becomes possible again, archaeological grant agencies must fund studies of materials in museum collections and encourage desk-based research in collaboration with Libyan scientists.

MELTING POT

Stretching from the Mediterranean Sea to the heart of the Sahara, Libya was a crossroads for many ancient cultures, including the Phoenicians, Greeks and Romans. It hosts five UNESCO World Heritage Sites that illustrate the country’s historical diversity (see ‘Threatened heritage’): Cyrene, founded in about 630 BC by the Greeks, was a principal town of the Hellenic world;

Leptis Magna, once part of the Phoenician city-state of Carthage, was incorporated in 46 BC by the Romans into the province of Africa; Sabraha, west of modern Tripoli, was a Phoenician trading post that became an influential Roman town during the second and third centuries AD; the old town of Ghatames, known as the ‘pearl of the desert’, is noted for its outstanding traditional architecture and was earlier home to the Romans and Berbers; and the Tadrart Acacus mountains are rich in prehistoric rock art.

These fragile vestiges of the human past are vulnerable to natural and human threats. Harsh environmental conditions — temperature variations and wind erosion — take a toll on rock art and open-air archaeological sites in the desert. These sites are threatened by infrastructure development, reclamation of land for agriculture, exploitation of underground resources such as oil, water and gas, and vandalism. The same applies to sites in towns and villages, such as the classical cities along the Mediterranean coast and the late-first-millennium BC Garamantian cemetery in the ancient Wadi al-Ajal river valley near the town of Germa.

DYING DISCIPLINE

The Tadrart Acacus is a place of unbelievable beauty; it used to be a global tourist destination. Some of its sites were vandalized in 2009; further damage — mostly graffiti — has been reported.

Today, the site is inaccessible: no commercial flight connects Tripoli and Ghat, a nearby town (a weekly military aircraft

brings food, essential goods and first-aid equipment). The tarred road between Ghat and Ubari is broken up, and clashes between the Tebu and Tuareg tribes increasingly affect the area.

Perhaps the greatest threat to Libya’s diverse heritage is the trafficking of archaeological materials, for profit or to fund radical groups. This has already been documented in Syria and Iraq⁷. No one has been able to fully assess the situation in Libya. Going to work among the black smoke of grenades, the men and women of the Libyan Department of Antiquities are doing their best. But museums are closed and the little activity left in the field is limited to the north.

LOST OPPORTUNITY

The Gaddafi regime neglected Libyan prehistory and relegated it to folklore. Classical towns such as Sabraha, Cyrene and Leptis Magna were viewed negatively as links to a colonialist past.

Among the hopes sparked by the revolution was the idea of a more modern view of the archaeological and cultural heritage — as a gateway to a shared national identity, a major revenue source and a focus for forging relationships with the rest of the world. Those hopes have been dashed.

The international community took actions to safeguard Libyan heritage at a UNESCO meeting in Paris at the end of October 2011, while the revolutionaries were still fighting. I learned, in a small Parisian café, that Gaddafi had been killed. The next day, hopes for Libya’s future filled the meeting room. Participants unanimously decided to build a shared programme for the training of



Ghadames (left), previously home to the Romans, suffered a rocket attack in 2011; tenth-century Islamic tombs at Zuwila (right) were damaged last year.

young Libyan archaeologists and scientists in Cyrenaica — the region where the revolution started — and Shahhat (modern Cyrene).

Partnerships between the Libyan government and UNESCO, with funding from Italy, were forged to strengthen Libya's research and stewardship capacity in archaeology. Particular attention was given to building the national archive and training Libyan police, customs officers and workers in the antiquities department in the fight against the trafficking of cultural property.

But few initiatives reached the field and now all efforts have essentially stopped. The escalation of hostilities in the past six months has led many foreign embassies to close and international organizations to move to neighbouring Tunisia. Large parts of North Africa are cut off by the civil war in Libya, the Tuareg rebellion in Niger, insecurity in Chad, Algeria and Mali, the presence of al-Qaeda and ISIS militants, and a large and uncontrolled circulation of weapons. International granting bodies such as universities and European and US research agencies have ceased to support field expeditions.

Being a Saharan archaeologist today is a difficult job. Researchers fear being kidnapped or even killed. Insurance cover is hard to come by. We could transfer our activities to another place, but this would mean abandoning decades of ideas, investment, and relations built with friends and colleagues. It would be hypocritical and sad.

I strongly believe that scientific cooperation is an effective way to bring people closer, increase confidence and make cultures more open to one other.

REKINDLE RESEARCH

Fieldwork is vital to research and central to fundraising in archaeology. But in Libya — and other violence-racked countries — archaeology as we have practised it has come to an end. Lengthy excavation campaigns will be impossible for years, if not generations. Researchers must imagine a different future based on other methods.

International funding and attention must return to scientific studies of Libyan heritage. Research should focus on existing materials in museums and collections. Granting bodies should give greater priority to research that can be carried out on computers or in the laboratory. Sample analyses of archaeological materials can be done in international labs, where Libyan scientists should work and be trained.

Building an online library of rock-art sites, with the involvement of Libyan students and colleagues from other countries, would help Libyan scientists to overcome their isolation and regain a sense of identity. Museum collections that span from remote prehistory to the Islamic cultures should be digitized and made freely available to a global audience. Unpublished collections held by international teams should also be digitized and shared online. Remote analyses of satellite imagery, for example, has been used to reveal lost Saharan cities (see go.nature.com/8y1gxh).

International cooperation between local

and foreign groups working in Libya must be supported. Travel funding and visas for Libyan scientists to work temporarily overseas should be found. And mobility programmes for scientists such as the European Union's Erasmus Mundus should be exploited — Libya's application numbers have been historically low. Energy companies and others with commercial interests in Libya should be encouraged to work with local stakeholders to help to train local personnel in scientific research.

Without these steps, archaeological research in Libya, already moribund, will soon die. It would be gravely disappointing and paradoxical if after years of neglect under the Gaddafi regime Libyan archaeological heritage is once again abandoned. As well as a failure of the 2011 revolution, it would be a missed opportunity for a generation of young Libyan archaeologists — and a tragedy for the safeguarding of monuments and sites of universal and outstanding value. ■

Savino di Lernia is director of *The Archaeological Mission in the Sahara*, Sapienza University of Rome, Italy.
e-mail: savino.dilernia@uniroma1.it

1. Drake, N. A., Blench, R. M., Armitage, S. J., Bristow, C. S. & White, K. H. *Proc. Natl Acad. Sci. USA* **108**, 458–462 (2011).
2. deMenocal, P. B. & Tierney, J. E. *Nature Edu. Knowl.* **3**, 12 (2012).
3. Kuper, R. & Kröppel, S. *Science* **313**, 803–807 (2006).
4. Balter, M. *Science* **331**, 20–23 (2011).
5. Dunne, J. *et al. Nature* **486**, 390–394 (2012).
6. Mattingly, D. J. & Sterry, M. *Antiquity* **87**, 503–518 (2013).
7. Lawler, A. *Science* **346**, 1162–1163 (2014).



A US soldier assists a member of his unit during the Battle of Okinawa in the Second World War.

W. EUGENE SMITH/MAGNUM PHOTOS

SOCIOBIOLOGY

Altruists together

Herbert Gintis applauds two books that powerfully enrich the dialogue on behavioural science.

Are humans basically selfish yet browbeaten by society into curbing their instincts? Or are they basically altruistic but corrupted by unjust societies? These age-old questions are now asked by behavioural scientists and discussed in journals such as *Nature*. Evolutionary biologist David Sloan Wilson's *Does Altruism Exist?* and science historian Michael Shermer's *The Moral Arc* are brilliant contributions to this branch of sociopolitical discourse.

Applying scientific principles to human society is hard. Society is a complex dynamical, adaptive nonlinear system. Moreover, rapid technical change, increased population density and globalization mean that we cannot reliably predict the future from the past. Even human nature, forged tens of thousands of years ago, turns out to be stunningly plastic.

Wilson's question is: do actions that mainly benefit unrelated others at personal cost exist? Could anyone doubt it? We give to charity,

Does Altruism Exist?: Culture, Genes, and the Welfare of Others

DAVID SLOAN WILSON
Yale University Press: 2015.

The Moral Arc: How Science and Reason Lead Humanity toward Truth, Justice, and Freedom

MICHAEL SHERMER
Henry Holt: 2015.

vote for public education even when we have no children, and volunteer to fight and die in war. People conform to social norms even when no one is looking, and punish the anti-social behaviour of others even when it is costly to do so. Yet for decades, a countervailing theory has held in biology and economics.

Richard Dawkins, in *The Selfish Gene* (Oxford University Press, 1976), reflected the opinion then current among biologists: "Let us try to teach generosity and altruism, because we are born selfish." Some 35 years later, in *Nature*, 137 evolutionary biologists

petitioned that "natural selection leads organisms to become adapted as if to maximize their inclusive fitness" (P. Abbot *et al. Nature* 471, E1–E4; 2011) — even in the most highly social species, individuals primarily help relatives. In fact, inclusive-fitness maximization is a pious wish of many population biologists that has never been validated in theory or fact.

Wilson's basic principle is the group-selection credo: "Selfishness beats altruism within groups. Altruistic groups beat selfish groups. Everything else is commentary" (D. S. Wilson and E. O. Wilson *Q. Rev. Biol.* 82, 327–348; 2007). As Charles Darwin noted in *The Descent of Man* (Murray, 1871), a hunter-gatherer band with many brave, altruistic soldiers will triumph over a group made up mostly of selfish cowards, even though the best thing of all for an individual is to be a coward surrounded by brave compatriots. The mathematics supports this scenario.

It is fashionable to question this view, but the theoretical issues have been resolved for decades. Groups do not mate or produce offspring, and so do not have biological fitness. Rather, the social organization of a species, its mating patterns and social groupings, is inscribed in the genomes of species members. Groups with successful social organization tend to enhance the fitness of their members, whose genomes code for this organization. Altruism can evolve in such groups, provided that altruists tend to be grouped preferentially with other altruists, in which case their biological fitness can on average be at least as high as that of selfish types.

As Wilson shows, another important source of human success is that cultures stress cooperation within the group, and so punish antagonistic individuals. This has led to humans 'domesticating themselves', favouring a human nature that is relatively docile and dependent on the company and approval of others. Moreover, humans have evolved to coordinate their behaviour, each member of a team 'reading the minds' of the others and identifying with common goals (see Michael Tomasello's *A Natural History of Human Thinking*; Harvard University Press, 2014).

Shermer's *The Moral Arc*, although grounded in behavioural game theory and social psychology, is the more speculative book. He offers a defence of science and reason as emancipatory tools in the face of bigotry, pseudoscience and faith. He, too, argues that humans are basically moral and cooperative, but adds that they are parochial. When their community is threatened, people turn compassion for kin into hatred for outsiders.

Shermer's central point is that even evil people are generally motivated by their own particular morals. In the perpetrators' minds, violence against outsiders is the application of justice. This requires that the enemy be deemed inferior and the cause of problems — an excuse historically manipulated by

Machiavellian leaders to gather support for their ambitions, as when the Nazis blamed the Jewish people for Germany's economic woes.

This is where science, technology and reason come into play, Shermer argues: the growth of global information and communications networks has rendered it increasingly difficult to perpetrate the falsehoods that let authoritarian leaders maintain their rule. For Shermer, an increasingly educated populace with access to information undermines parochialism and pseudoscience, by allowing people to judge for themselves. The role of smartphones and social media in fuelling the 2011 Arab Spring uprisings is a case in point.

This is a welcome turnaround from *The Believing Brain* (Times, 2011), in which Shermer argued the rather nihilistic position that “beliefs come first, explanations for beliefs follow”. In *The Moral Arc*, Shermer, founder of the Skeptics Society, adheres to Enlightenment thought. His subtitle, *How Science and Reason Lead Humanity Toward Truth, Justice, and Freedom*, evokes the call to arms of philosopher Immanuel Kant in his 1784 *What is Enlightenment?*: “Have the courage to use your own understanding.”

Some of Shermer's positions would have surprised Enlightenment writers. Kant, for instance, believed that the oppressive state and authoritarian church were the sole impediments to truth and justice. We know now that even people with access to the ballot box and free expression can embrace intolerant and obscurantist doctrines. Moreover, Voltaire and others believed that the uneducated could not apply reason to the affairs of life. Shermer, by contrast, is a vigorous proponent of political democracy and equal rights.

Shermer's is an exciting vision, but he is mistaken in thinking that truth, freedom and justice are the inevitable by-products of scientific advance. Modern liberal democracy is the product of masses of people collectively throwing off the yoke of authoritarian states. But the power of popular action was made possible by a military technology: the handgun. This displaced elite cavalry and required nations to give the vote to peasants and citizens, who became the lifeblood of military defence. Even today, the United States, with its formidable drones and missiles, cannot win a war without ‘troops on the ground’.

We must be on constant guard against new instruments of information control, persecution and death that could once again render secular and religious totalitarianism a viable social alternative. Constant vigilance by altruists such as Wilson and rationalists such as Shermer may in the end win the day. ■

Herbert Gintis is external professor at the Santa Fe Institute in New Mexico. His most recent book is *A Cooperative Species* (with Samuel Bowles).
e-mail: hgintis@comcast.net

Books in brief



Most Wanted Particle: The Inside Story of the Hunt for the Higgs, the Heart of the Future of Physics

Jon Butterworth EXPERIMENT (2015)

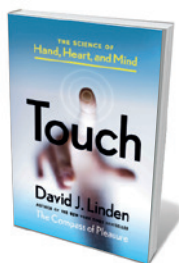
The Higgs boson may seem amply biographized, but Jon Butterworth's account of its 2012 discovery offers deep context. As a physicist on the ATLAS experiment at CERN — the Higgs hunting ground near Geneva, Switzerland — Butterworth is an insider's insider. His narrative seethes with insights on the project's science, technology and 'tribes', as well as his personal (and often amusing) journey as a frontier physicist. Glossaries on the standard model of physics, Feynman diagrams and more are included.



Sea of Storms: A History of Hurricanes in the Greater Caribbean from Columbus to Katrina

Stuart B. Schwartz PRINCETON UNIVERSITY PRESS (2015)

Ten years ago, Hurricane Katrina killed more than 1,800 people and submerged 80% of New Orleans, Louisiana. Historian Stuart Schwartz frames that catastrophe within five centuries of hurricanes in the greater Caribbean — natural disasters that mirrored and exacerbated the violent social upheavals that erupted as European nations pursued New World riches. Today, a mix of political vagaries and patchy official disaster response presents dangerous ambiguities in a region where more cyclones are a certainty.



Touch: The Science of Hand, Heart, and Mind

David J. Linden VIKING ADULT (2015)

A touching story? A tactless comment? So elemental is the sense of touch that it permeates metaphors we live by. In this succinct treatise, neuroscientist David Linden explores the “weird, complex, and often counter-intuitive” tactile system and its intimate impact on the human experience. Through scores of scientific studies and anecdotes, Linden investigates phenomena ranging from the two separate touch systems in the skin (one slow, one fast), to a detailed ‘cast list’ for the main neurophysiological players in orgasm, such as the somatosensory cortex, amygdala and cerebellar nuclei.



Melting Away: A Ten-Year Journey through Our Endangered Polar Regions

Camille Seaman PRINCETON ARCHITECTURAL PRESS (2014)

In the space of a generation, Antarctica and the Arctic have metamorphosed from remote frontiers to cruise destinations, their icy reaches and charismatic wildlife exhaustively mapped and filmed. But writer and photographer Camille Seaman (see J. Hoffman *Nature* **492**, 40; 2012) has a rare gift for making them seem arrestingly alien again. Her coffee-table book is the product of ten years at the poles; its images alone are a compelling argument for protecting the wonder and strangeness at the ends of the Earth.



Fantasy Islands: Chinese Dreams and Ecological Fears in an Age of Climate Crisis

Julie Sze UNIVERSITY OF CALIFORNIA PRESS (2015)

Carbon-neutral, zero-waste and home to 500,000 people: the Chinese eco-city of Dongtan seemed a radical urban dream. But the city, to be sited near Shanghai on Chongming — the world's biggest alluvial island — remains a blueprint. As Julie Sze argues in this thoughtful, if uneven, analysis of Chinese “eco-desire”, the culprit could be irreconcilable beliefs in harmony with nature, and the ability of autocratic political structures to enact radical change. *Barbara Kiser*

Correspondence

In praise of Holt as head of the AAAS

The American Association for the Advancement of Science (AAAS) has chosen Rush Holt, who was a Democratic congressman for eight terms, as its new chief executive. Daniel Sarewitz attacks this choice as “political” (*Nature* **516**, 9; 2014), but it is not partisan.

The AAAS announcement praises Holt for broadly promoting “the value of science communication, particularly for conveying information about climate change”. In its March 2014 Climate Science Panel report, the AAAS talked bluntly about the dangers of inaction and of poor science communication — a view you share in calling on scientists to ensure that “they are not bested in the court of public opinion” (*Nature* **464**, 141; 2010).

Sarewitz contends that the AAAS is “anointing a leader who could take up the fight” with climate-science deniers, among whom are many Republican politicians. But in Holt, the AAAS has a scientist who understands the fight that we are in (see also *Nature* **471**, 265–266; 2011) and is well placed to defend it from attacks by Congress. He is an inspired choice.

Joseph Romm Center for American Progress, Washington DC, USA.
jromm@americanprogress.org

Shale gas: nuance in output predictions

You claim that the most recent estimates of future output for shale gas in the United States have become more conservative, but in our view this is a red herring (*Nature* **516**, 28–30; 2014).

State-of-the-art projections for the world’s future shale-gas supplies hinge on improved quantification of the uncertainty range and reducing its spread (called ‘de-risking’ a shale play) as experience and technology advance. Several of the studies you quote include uncertainty

ranges that explain the current spread in forward projections of future US gas supplies. That crucial nuance was missing from your graphic, however, which shows only a simplified, discrete forward-production prognosis.

Comparing just one scenario from the study by the team at the University of Texas at Austin with another from the US Energy Information Administration’s shale-gas outlook, omitting uncertainty ranges, creates an apparent mismatch where one may not in fact exist.

As a result of technology innovation (see also *Nature* **516**, 7; 2014), the United States is today drilling 3-kilometre-long horizontal wells and conducting 30-stage fracture treatments at depths of 3.7 km. Further technological gains will increase global oil and gas output (see, for example, S. Neff and M. Coleman *Energy Strategy Rev.* **5**, 6–13; 2014). Oil and gas prices also drive global shale development.

No one can accurately predict both the technology improvement rate and future wellhead prices, so we have to rely on a range of forecasts based on a variety of assumptions.

Steve Holditch, Dan Hill, Ruud Weijermars Texas A&M University, College Station, USA.
r.weijermars@pe.tamu.edu

Shale gas: hardly a fallacy

We believe that your comparison of US fracking forecasts creates a false dichotomy between modelling results from the US Energy Information Administration (EIA) and the Bureau of Economic Geology at the University of Texas at Austin (*Nature* **516**, 28–30; 2014).

Our integrated team of scientists, engineers and economists at the University of Texas has built rigorous models that incorporate a wide range of input variables and well-constrained outcome scenarios. In our view, the comparison of

just one simulation run with a single outlook from the EIA trivializes a complex problem and fails to represent accurately the rigour and uniqueness of what is being accomplished in our four-year study (see go.nature.com/zfverj).

Your graphic ‘Battle of the forecasts’ is partially attributed to our data. Although we present preliminary results at conferences and make them available on our website, we explained to the author that our work on the Haynesville and Marcellus plays was not yet finished or published, and requested that it should not be used. We therefore question why you should choose to base the main thread of your argument on a comparison to our unfinished work.

Finally, I find your headline ‘The fracking fallacy’ potentially misleading: in isolation, it reads as a negative comment on the fracking process itself, rather than on forecasts of natural gas production. Production of US oil is currently at a 30-year high, and of natural gas at an all-time high. Hydraulically fractured wells account for more than half and almost half, respectively, of US natural gas and oil production. To imply otherwise does a disservice to your readers.

Scott W. Tinker, Svetlana Ikonnikova The University of Texas at Austin, Texas, USA.
scott.tinker@beg.utexas.edu

Editorial note: Scott Tinker and Svetlana Ikonnikova informed *Nature* that their study was unpublished. They subsequently made the data publicly available, at which point *Nature* used that information and gave appropriate credit.

United Nations highlights soil crisis

Some 500 years after Leonardo da Vinci declared that more was known about celestial bodies than about the soil underfoot, the United Nations has proclaimed

2015 the International Year of Soils. This offers a unique opportunity to address the crisis in soil sustainability (see www.fao.org/soils-2015).

Among the factors undermining soil quality are intensive farming, industrial activity and increasing urbanization. Soil contamination is threatening food production, water potability and ecosystem services, notably in large parts of China. Safeguarding soils is therefore crucial to the UN Post-2015 Development Agenda and the Sustainable Development Goals (go.nature.com/s7jcik).

Initiatives that are already under way for the sustainable management of complex soil systems include the Intergovernmental Technical Panel on Soils, the Global Soil Biodiversity Initiative, the Global Network of Critical Zone Observatories, and the International Soil Modeling Consortium.

Henry Lin Institute of Earth Environment, Chinese Academy of Sciences, Xi’an, China; and Pennsylvania State University, University Park, USA.

Rainer Horn Christian Albrechts University zu Kiel, Germany.
henrylin@psu.edu

CORRECTIONS

The ‘West Asia’ article in the *Nature Index* (*Nature* **515**, S88–S89; 2014) stated that King Abdullah University of Science and Technology had an article count of 121 and a weighted fractional count of 9.96. In fact, it was King Abdulaziz University that had these values.

In the *Nature Index China*, the ‘Chinese Academy of Sciences’ article (*Nature* **516**, S56–S57; 2014) should have affiliated Peng Zhang to the Institute of Plant Physiology and Ecology. And in the ‘Beijing’ article (*Nature* **516**, S60–S61; 2014), Ning Jiao’s quote was mistranslated, so it has been updated.

Donald Metcalf

(1929–2014)

Discoverer of hormones that regulate blood-cell proliferation.

Donald Metcalf established which blood cells give rise to which, and identified the hormones that regulate the cells' proliferation and differentiation. His work, which shed light on how to boost people's supplies of white blood cells, has benefited millions.

Metcalf — Don to nearly everyone who worked with him — died on 15 December 2014. He was born in 1929 in Mittagong in the Southern Highlands of New South Wales, Australia, to schoolteachers. During his medical degree at the University of Sydney, Metcalf's passion for research was ignited when he spent a year studying the ectromelia virus, which causes 'mousepox'.

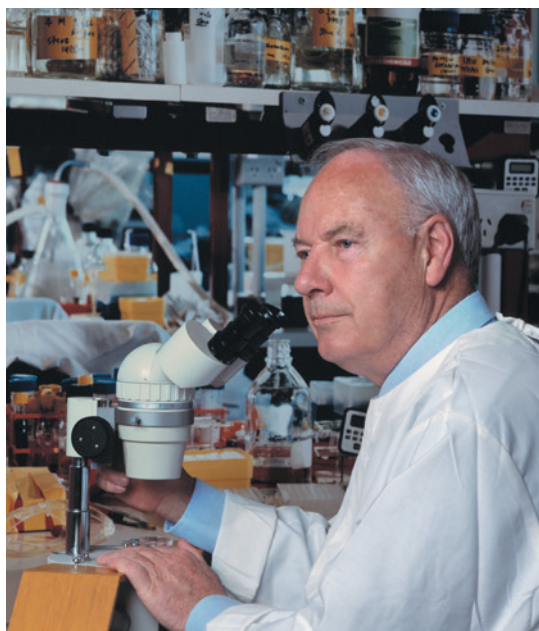
Metcalf received his degree in 1953 and moved to Melbourne in 1954 to join the Walter and Eliza Hall Institute of Medical Research (WEHI). Aside from brief trips to Europe and the United States, he spent his 60-year career at the institute, supported throughout by the Carden Fellowship of the Anti-Cancer Council of Victoria (now the Cancer Council).

Metcalf's arrival at the WEHI was not smooth. Its then director, Frank Macfarlane Burnet, who was later awarded the Nobel Prize in Physiology or Medicine for his work on immunity, was not a fan of cancer research. According to Metcalf, Burnet, like many others at the time, viewed cancer as "an inevitable disease", and cancer researchers as either "rogues or fools". Metcalf was undeterred.

For ten years he worked on cell turnover in an immune organ called the thymus, until a chance finding changed his focus. Metcalf, with Ray Bradley of the University of Melbourne, discovered that he could grow colonies of blood cells in agar, provided that the right stimulus was added.

Metcalf's genius lay in realizing that he could use this system to work out how cell types are related, and also to characterize the hormones that regulate the cells' proliferation and differentiation. He named these hormones colony-stimulating factors (CSFs). Over the next 50 years, Metcalf made the blood-cell system the model for understanding the regulation of cell growth in body tissue.

Early on, Metcalf realized that to purify CSFs, and to clone the genes that encoded them, he would need collaborators. He



recruited young faculty members with the required biochemical and biophysical skills to work with him at the WEHI. The group later collaborated with several molecular biologists who came to work at the Melbourne branch of the international Ludwig Institute for Cancer Research. The branch was directed by Tony Burgess, previously a laboratory head in Metcalf's WEHI Cancer Research Unit.

From 1965 to 1985, Metcalf and his team identified and purified four colony-stimulating factors: granulocyte-macrophage CSF (GM-CSF), granulocyte CSF (G-CSF), macrophage CSF (M-CSF) and multi-CSF, now known as interleukin-3. The team also cloned the gene for one of these, GM-CSF; the other genes were cloned by groups around the world.

This work paved the way for mass production of the hormones and the experiment that Metcalf had long dreamed of: injecting CSFs into animals. His nagging doubt was that the factors — purified using a contrived *in vitro* assay — might be irrelevant to normal physiology. He needn't have worried. In mice, the CSFs triggered a spectacular rise in the number of white blood cells in the bone marrow and peripheral blood.

Clinical applications soon followed. The most widespread use of CSFs has been to ameliorate leukopenia, a decline in the number of white blood cells associated with

chemotherapy. During phase I clinical trials of G-CSF conducted in the late 1980s, Metcalf and his collaborators noticed that administering the hormone prompted large numbers of haemopoietic stem cells — precursor cells that give rise to all types of blood cell — to move from a person's bone marrow into their peripheral blood.

The finding allowed clinicians to harvest stem cells simply by injecting people with G-CSF and taking their blood, rather than by extracting the cells from bone marrow — a more painful and complicated procedure. The new method made the transplantation of blood stem cells safer, easier, more effective and ultimately more widely used. In the past 20 years, 20 million people are thought to have benefitted from Metcalf's discoveries.

None of the many prizes that Don received conveys the degree to which he was a scientist's scientist. He distrusted researchers who had turned their back on the bench; he always worked in the laboratory, assisted by one or two research assistants and an occasional graduate student or postdoctoral fellow. He had an incredible work ethic. Having worked in the lab for eight or nine hours, Don would write papers or books at home. He also detested spin. His inclination was to produce one paragraph of discussion for each page of results — not a word more.

The only thing that Don valued more than his science was his family — Jo, his wife of more than 60 years, his four daughters and six grandchildren. Last August, when Don was diagnosed with incurable metastatic pancreatic cancer, he faced a dilemma: how could he continue to do experiments and spend as much time as possible with his beloved Jo? He found a solution: he had his microscope moved to his dining-room table. Don continued to work, surrounded by his loved ones, until early November — exactly as he wanted it. ■

Douglas Hilton is director of the Walter and Eliza Hall Institute of Medical Research in Melbourne, Australia, and head of the Department of Medical Biology at the University of Melbourne. He first worked with Don Metcalf as an undergraduate in the 1980s.
e-mail: hilton@wehi.edu.au

WALTER AND ELIZA HALL INSTITUTE

STEM CELLS

Emergency back-up for lung repair

Influenza virus severely damages the epithelial tissue that lines the lung. Findings suggest that, in mice, activation of a back-up population of stem cells mediates effective repair of the injured lung. [SEE LETTERS P.616 & P.621](#)

EMMA L. RAWLINS

The lung has low rates of day-to-day cell turnover, but a tremendous potential for repair following injury. Despite this regenerative ability, the existence of dedicated lung stem cells has been hotly contested¹, and the characteristics of any stem-cell population that may reside in this tissue are unknown. In two papers in this issue, Vaughan *et al.*² (page 621) and Zuo *et al.*³ (page 616) show that when the epithelial cells lining the interior of the lung are damaged by infection with influenza virus, a rare stem-cell population is induced to proliferate and migrate to the damaged site. There, this population can differentiate into several cell types.

Over the past five years, it has become clear that many types of the differentiated secretory epithelial cells in the lung can function as stem cells at steady state and following mild injury^{4–6}. But what happens in response to severe injury is less clear. Infecting mice with influenza is a useful way to severely injure lung cells, both to investigate cell behaviour and to model the human disease⁷. In response to influenza, a cell population seems to migrate from the airways into the alveoli, where they remain⁷. However, the identity of these cells and their specific contribution to alveolar repair have been unknown.

The current papers address this issue using lineage tracing in mice, in which a genetic trick indelibly marks a cell population of interest and all of its descendants. Vaughan and colleagues traced populations of differentiated secretory epithelial cells, which maintain normal lungs and repair minor injuries^{4–6} (Fig. 1a), and showed that these cells are not involved in influenza-induced repair. Importantly, they also showed that spurious results could be obtained if care was not taken with the lineage-tracing technique. Both groups defined a previously unnoticed cell population containing stem cells that mediate repair following influenza.

The stem cells, which are located in the airways, were dubbed distal airway stem cells by Zuo *et al.*, and lineage-negative progenitors by Vaughan and co-workers. They are rare, undifferentiated, basally located cells — that is, their top surface does not extend

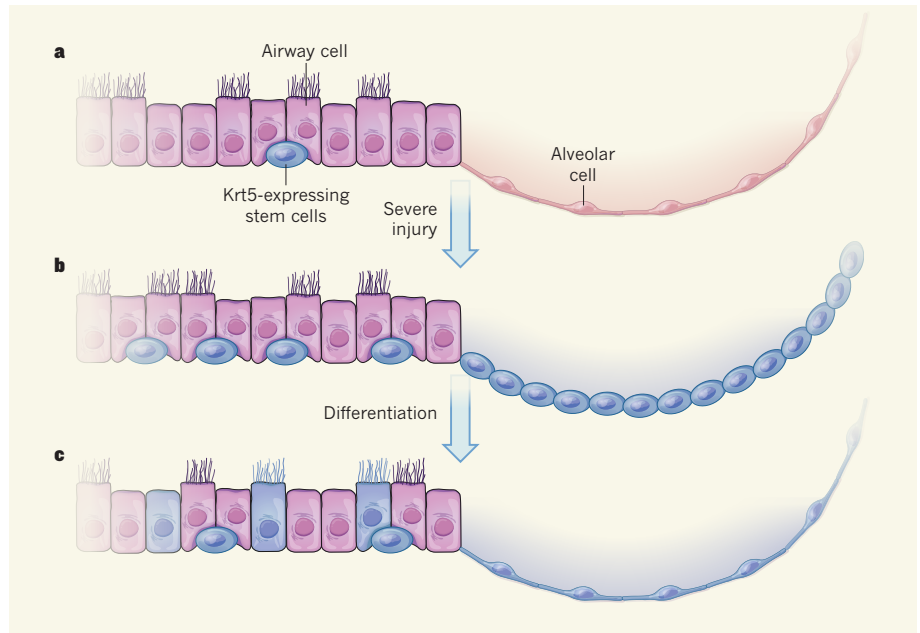


Figure 1 | Stem cells that mediate lung repair. **a**, The airways and alveoli of the lung are lined with distinctive epithelial cells. Vaughan *et al.*² and Zuo *et al.*³ identify a population of Krt5-expressing stem cells that are rare and inactive under steady-state or mild-injury conditions. **b**, Krt5-expressing stem cells are activated in response to the severe lung damage caused by influenza infection. These cells proliferate and migrate into the alveoli. **c**, Zuo and colleagues report that these cells differentiate into both airway and alveolar lineages, resulting in productive lung repair. By contrast, in the mice analysed by Vaughan and co-workers, the cells differentiated into airway lineages, but remained undifferentiated in the alveolar epithelium until experimental inhibition of the Notch signalling pathway allowed differentiation to occur.

up to the air space of the airway. The cells express the cytokeratin 5 (Krt5) protein and the transcription factor Trp63, either singly or together. These molecules are widely expressed in basally located, dedicated stem cells in other epithelial tissues⁸.

Both groups traced a Krt5-expressing subset of the cells and showed that these cells move into the alveoli following injury (Fig. 1b). Furthermore, Vaughan *et al.* used time-lapse microscopy to analyse lung slices through live-cell imaging. They found that the Krt5-expressing stem cells can migrate through alveolar walls, indicating that epithelial cells can migrate farther to contribute to repair than previously thought. By contrast, Zuo and colleagues killed the cells at the beginning of the migration process. This resulted in failure of alveolar repair, thus demonstrating that these stem cells are necessary for effective repair.

Despite this broad agreement between the

two studies, they differ in interesting ways. The groups traced Krt5-expressing stem cells and their descendants using mice that carried slightly different genetic modifications^{8,9}. Zuo *et al.* demonstrated that the cells they traced differentiate into mature alveolar cells (Fig. 1c). By contrast, the cells traced by Vaughan *et al.* move into the alveoli and remain there after infection, but never differentiate into mature cells.

How do we reconcile these differences? One possibility is a discrepancy in the extent of injury. The route by which each group infected their mice was different (intratracheal or intranasal). There may also have been small differences in virus levels, or in the genetic make-up of the mice used. To avoid confusion in the future, influenza models should be standardized, or the extent of injury characterized in each experiment. Alternatively, the two laboratories may have been sampling slightly different subsets of cells from the same population.

It remains unclear how many stem-cell populations there are in the lung, although the Krt5-expressing population characterized in these studies seems to reconcile previous, apparently disparate, findings^{7,10}. But how many cell types are included in this population? Vaughan and colleagues analysed the transcriptional profiles of individual stem cells, and provided preliminary evidence that the population contains many different cell types (it is heterogeneous). However, this may reflect current limitations in techniques for isolating pure populations of these cells. A comprehensive answer to the question will require *in vivo* experiments in which individual stem cells are lineage-traced, coupled with further single-cell molecular analysis. Such experiments would determine whether the cells constitute a single population, and so the different results reflect different experimental conditions, or whether the population is truly heterogeneous and perhaps contains cells with different capacities for migration or differentiation.

Nevertheless, the differences between the two studies mean that each provides distinct insights that could be relevant to human health. Vaughan and co-workers found that chemical inhibition of the Notch signalling pathway was required for their cells to differentiate into mature alveolar cells (Fig. 1c). Moreover, they identified regions of hyperactive Notch signalling in human lungs that had defective alveolar repair, implicating this pathway in the development of chronic lung conditions. By contrast, Zuo *et al.* reported that Krt5-expressing stem cells could be grown in culture and transplanted into influenza-infected mice lacking their own stem cells. This restored the repair process, opening up the possibility that stem-cell therapy will eventually be used to treat lung conditions. These findings require much more fundamental research. But there are many lung conditions for which only palliative therapies are available — a powerful incentive to explore the possibilities.

A key question is how these lung stem cells interact with their associated mesenchymal (non-epithelial) cells to mediate a functional repair process. In the adult skin, mesenchymal cell populations from differing embryonic origins have distinct roles in wound repair¹¹. Lung mesenchymal cell populations are just beginning to be characterized¹². If any medical applications are to arise from studies of lung repair, the full cellular picture will be required. ■

Emma L. Rawlins is at the Wellcome Trust/Cancer Research UK Gurdon Institute, the Wellcome Trust/MRC Stem Cell Institute and in the Department of Pathology, University of Cambridge, Cambridge CB2 1QN, UK.
e-mail: e.rawlins@gurdon.cam.ac.uk

1. Hogan, B. L. M. *et al.* *Cell Stem Cell* **15**, 123–138 (2014).
2. Vaughan, A. E. *et al.* *Nature* **517**, 621–625 (2015).
3. Zuo, W. *et al.* *Nature* **517**, 616–620 (2015).
4. Barkauskas, C. E. *et al.* *J. Clin. Invest.*

123, 3025–3036 (2013).

5. Desai, T. J., Brownfield, D. G. & Krasnow, M. A. *Nature* **507**, 190–194 (2014).
6. Rawlins, E. L. *et al.* *Cell Stem Cell* **4**, 525–534 (2009).
7. Kumar, P. A. *et al.* *Cell* **147**, 525–538 (2011).
8. Rock, J. R. *et al.* *Proc. Natl Acad. Sci. USA* **106**,

12771–12775 (2009).

9. Van Keymeulen, A. *et al.* *Nature* **479**, 189–193 (2011).
10. Chapman, H. A. *et al.* *J. Clin. Invest.* **121**, 2855–2862 (2011).
11. Driskell, R. R. *et al.* *Nature* **504**, 277–281 (2013).
12. Kumar, M. E. *et al.* *Science* **346**, 6211 (2014).

ASTROPHYSICS

Stellar clocks

A link between rotation and age for Sun-like stars has long been known, but a stringent test of it for older stars has been lacking. The Kepler mission helps to fill this gap with observations of an old star cluster. SEE LETTER P.589

DAVID SODERBLOM

The clocks of the cosmos tick constantly, but too softly for us to hear, and so we cannot directly measure the ages of stars. Yet we need good ages, because nearly every aspect of astrophysics deals with how things evolve with time. Even the Sun is silent on this score, and it is only from being able to study Solar System material — meteorites — in the laboratory that we know the Sun's age so exquisitely: $4,567 \pm 1 \pm 5$ million years¹. In this issue, Meibom *et al.*² (page 589) describe a key step needed to obtain better estimates for the ages of cool, low-mass stars like the Sun.

Many of the stars for which we would like to know the ages are like the Sun. They live for 10 billion years or more, and so represent the entire age of our Galaxy's disk, which is where most of the Galactic stars reside. Sun-like stars are the ones that naturally excite the most interest in our quest for other Earths, and, indeed, the first question that will be asked when someone reports signs of life on an exoplanet will be, 'how old is the host star?', because we will want to place the discovery in an evolutionary context.

What makes the Sun and stars like it

favourable for life-bearing planets is that they change very slowly over time. But that also makes them inaccessible to age estimation by conventional means, which is by determining a star's temperature, luminosity and composition and then comparing them with stellar models. Those models themselves are calibrated against the Sun, the only star with well-established fundamental properties.

Given this problem of estimating ages, we settle for what we can get³, and it has been suspected for some time that, for Sun-like stars, rotation declines with age on something close to a power-law relation⁴ — a nice straight line in a log–log plot. This is convenient, in that our concern with accuracy roughly scales as the age itself, and if we could reliably derive ages that are good to 5% or better, that would improve our knowledge of stellar and Galactic processes substantially.

But why does the Sun spin so slowly, as much as 100 times more slowly than some young Sun-like stars, and is the spin rate a reliable clock? As was shown⁵ in 1967, slow rotation is a property shared by all Sun-like stars, and the mechanism underlying the phenomenon starts right where the outer layers of the stars begin to become convective. Thus, it is convection that

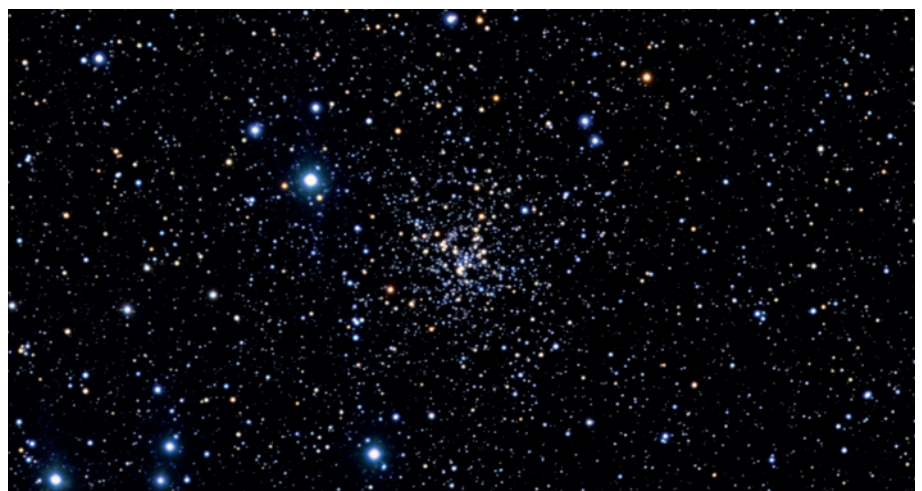


Figure 1 | Star cluster NGC 6819, the concentration of stars visible in the centre of the image.

makes the Sun and stars like it Sun-like. In the Sun, convection and rotation lead to complex motions of the conducting plasma of the convective outer layer. Those motions generate a magnetic dynamo that is based in or just below the outer layer. The magnetic field produced by the dynamo can then grip the solar wind of high-speed charged particles beyond the Sun's surface, and so angular momentum is steadily lost.

The loss of angular momentum is a slow process for the present-day Sun, but we also know that, because of their faster spin rate, young Suns generate much stronger magnetic fields. The magnetic dynamo provides a feedback mechanism that causes convergence in the spins of stars that have the same age but different initial rates. Observationally, that convergence seems to occur by about the age of the nearest loosely bound star cluster, the Hyades — that is, 600 million years or so.

Star clusters are fundamental to studies of angular-momentum loss because they provide good-sized samples of stars sharing the same composition and age. But the problem is that there are few clusters older than the Hyades. Clusters get ripped apart as they orbit in our Galaxy owing to tidal forces from objects such as giant molecular clouds or black holes (the source of the effect is still poorly understood). Such break-up accounts for why stars are spread all over the sky, as opposed to being concentrated in clusters, but it also means that clusters older than 0.5 billion years or so are inherently rare, and thus few are near the Sun. Compounding the difficulty, the faint stars in these more distant clusters have fewer and smaller star spots. These are regions of lower temperature than the surrounding surface that cause the star to dim and brighten when they rotate in and out of view, and so can be used to measure the star's spin. The few small spots on older stars yield variations in starlight that are hard to detect using ground-based telescopes.

That is why the capabilities of a mission such as NASA's Kepler satellite, with its exquisite measurements of stellar brightnesses, is essential, and so why Meibom and colleagues' work, which is based on Kepler data, matters. The authors used the satellite to measure the rotation period of 30 cool stars in the star cluster NGC 6819 (Fig. 1), which is about 2.5 billion years old. NGC 6819 therefore fills the large gap in age between the Sun and existing cluster observations. By using methods such as the study of rotation, Kepler has revolutionized stellar physics as much as it has the study of exoplanets. It means that we can consider using the slow, steady spin-down of Sun-like stars as a way of determining their ages.

But this method of estimating a star's age has limitations. The main one is that we do not understand the physics of rotation and angular-momentum loss in Sun-like stars, and the

rotation–age relation remains purely empirical. This is fine if we can calibrate that relation well and if there is a tight correspondence between rotation and age. But stars can acquire extra angular momentum late in their lives by swallowing a companion, such as another star or a planet. Orbiting objects have much more angular momentum than has a star, and so even small bodies can be significant. We know of no means by which a star can have its angular momentum stolen, and so these accumulations add a systematic uncertainty.

Also, as noted, older stars have at best only weak variations in their light. With even the highest-precision measurement, we cannot always see the signal of the Sun's rotation. Likewise, not all stars reveal their rotation to us even when we badly want them to; nature is indifferent to our curiosity. But persistence can

pry out those secrets. Meibom and colleagues' study shows exactly why we develop new capabilities: there is the good reason (in the case of Kepler, it was to find Earth-like planets, a very good reason), and then there is the real reason, which is to enable clever people to do what was not foreseen at the start. ■

David Soderblom is at the Space Telescope Science Institute, Baltimore, Maryland 21218, USA.
e-mail: drs@stsci.edu

1. Chaussidon, M. in *Lectures in Astrobiology: Volume II* (eds Gargaud, M., Martin, H. & Claeys, P.) 45–74 (Springer, 2007).
2. Meibom, S. *et al. Nature* **517**, 589–591 (2015).
3. Soderblom, D. R. *Annu. Rev. Astron. Astrophys.* **48**, 581–629 (2010).
4. Skumanich, A. *Astrophys. J.* **171**, 565–567 (1972).
5. Kraft, R. P. *Astrophys. J.* **150**, 551–570 (1967).

PLANT BIOLOGY

Seeing the wood and the trees

The identification of the gene regulatory network that controls the formation of xylem — the major component of wood — opens up new avenues for manipulating plant biomass. [SEE ARTICLE P.571](#)

ANTHONY BISHOPP & MALCOLM J. BENNETT

Cellulose, hemicelluloses and lignin are key natural polymers that make up the bulk of plant biomass¹. These biopolymers are also renewable resources for the production of dietary fibre, paper and biofuels². On page 571 of this issue, Taylor-Teeples *et al.*³ report the identification of the gene regulatory network that controls the synthesis of these biopolymers in root xylem cells of the model plant *Arabidopsis thaliana*.

Xylem is a plant tissue that provides mechanical support and the main mechanism for transporting water and nutrients from root to shoot tissues. To perform these important functions, xylem cells deposit a specially reinforced structure termed the secondary cell wall¹ (Fig. 1). Xylem secondary cell walls are composed mainly of cellulose, hemicelluloses and lignin. The cellulose forms a network of load-bearing fibres coated in hemicelluloses and embedded in lignin, providing mechanical strength and rigidity (akin to steel rods set in reinforced concrete). However, the presence of lignin is a major impediment to efficient extraction of the sugars in cellulose and hemicelluloses for their conversion to biofuels². Hence, understanding how the relative proportions of these biopolymers are controlled in plant tissue would open up opportunities

to redesign plants for biofuel use.

Xylem cells control the relative abundance of biopolymers in part by regulating expression of the genes that encode the enzymes for polymer synthesis⁴. Expression is controlled by transcription-factor proteins that bind DNA sequences, termed promoters, close to the genes. A handful of transcription factors have been identified that control the expression of individual genes regulating the production of cellulose, hemicelluloses and lignin. But this small-scale, gene-by-gene approach has provided a highly fragmented picture of the potential regulatory interactions between xylem-associated transcription factors and their gene targets.

Taylor-Teeples *et al.* instead adopted a network approach — screening more than 460 transcription factors expressed in the root xylem of *A. thaliana* for their ability to bind the promoters of around 50 previously characterized genes that encode cell-wall components or other transcription factors involved in xylem formation. This large-scale analysis provided a remarkable overview of the regulatory process, revealing a highly interconnected network composed of some 240 genes and more than 600 new protein–DNA interactions.

The xylem regulatory network shows that each cell-wall gene is bound, on average, by 5 different transcription factors, each

belonging to one of 35 distinct families of regulatory proteins. This regulatory arrangement provides a huge number of combinatorial possibilities, which Taylor-Teeples and colleagues show is crucial for integrating environmental signals such as salt or iron stress — alterations in the expression of certain transcription factors allowed different sub-networks to be used to adapt the cellular response to these conditions.

The network also reveals that many of the transcription factors are not part of simple linear pathways, but form a series of feed-forward loops (FFLs). Such regulatory systems are well recognized in systems biology, and typically involve a transcription factor that controls the expression of other transcription factors, which then collectively co-regulate their target genes. For example, the authors find that the transcription factor E2Fc binds to more than 20 promoters, including those for the genes encoding the transcription factors VND6, VND7 and MYB46, as well as genes associated with cellulose, hemicellulose and lignin production. Although FFLs are common in biological systems⁵, they are remarkably numerous in the xylem network, occurring close to 100 times. They are also frequently embedded within one another, creating FFL cascades. For example, the network shows that VND7 and MYB46 also bind to the promoters of many E2Fc target genes.

So why are there so many FFLs? Not only are there many possible components, but even for simple systems with only three components, there are many possible ways to wire them⁶. Common to all FFLs is a direct path (in which a source transcription factor regulates a target gene) and an indirect path (in which the same source factor regulates an intermediate transcription factor that regulates the same target). For 'coherent' FFLs, the direct and indirect paths have the same overall effect on the target gene (both activate or repress its expression), whereas for incoherent FFLs, one path activates and the other represses. Mathematical modelling of these loops has revealed that different arrangements can produce a range of responses from target genes. For example, coherent FFLs can protect against unwanted responses to fluctuations in inputs, whereas incoherent FFLs can speed up transcriptional responses⁶. In the case of the xylem network, a coherent FFL could result in tight regulation of cell-wall gene expression, thereby promoting secondary cell-wall synthesis in a switch-like manner to prevent the deposition of secondary cell-wall material in non-xylem cells.

It is not yet possible to determine exactly what types of FFL are present in the xylem regulatory network described by Taylor-Teeples *et al.*, because although the technology used by the authors identifies interactive nodes, it cannot predict whether they relate

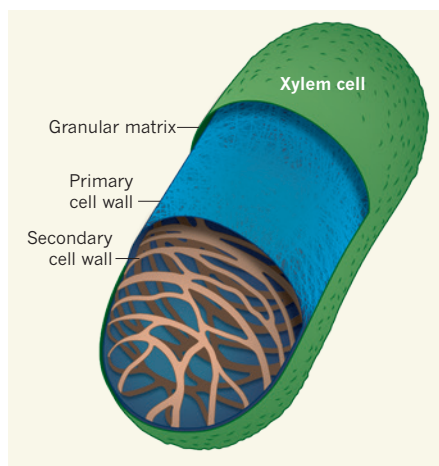


Figure 1 | Building a secondary cell wall. The cell wall of plant xylem cells (a tracheary-element cell grown *in vitro* is shown) contains several layers, including an outer granular matrix, a primary cell wall composed mainly of the biopolymer cellulose and an inner secondary cell wall composed of microfibrils of cellulose and lignin. Taylor-Teeples *et al.*³ have characterized the gene regulatory network that determines the production of xylem biopolymers. (Figure adapted from ref. 7.)

to transcriptional activation or repression. However, these nodes provide a framework for future research to characterize key interactions in a targeted, gene-by-gene manner, and to determine the precise regulatory

structure. This will allow the identification of ways to manipulate this network to engineer different cellular properties and develop new plant varieties for biofuel use. The description of the network also helps to explain why plant transcription factors have so far largely eluded identification by genetic screens, owing to functional redundancy among regulators of secondary-cell-wall biosynthesis. This knowledge can now be used to perform more-precisely targeted screens of gene function, by creating combinations of mutations that overcome this genetic redundancy. ■

Anthony Bishopp and Malcolm J. Bennett are in the Centre for Plant Integrative Biology, University of Nottingham, Sutton Bonington LE12 5RD, UK.

e-mails: anthony.bishopp@nottingham.ac.uk; malcolm.bennett@nottingham.ac.uk

1. Turner, S., Gallois, P. & Brown, D. *Annu. Rev. Plant Biol.* **58**, 407–433 (2007).
2. Somerville, C., Youngs, H., Taylor, C., Davis, S. C. & Long, S. P. *Science* **329**, 790–792 (2010).
3. Taylor-Teeples, M. *et al. Nature* **517**, 571–575 (2015).
4. Kondo, Y., Tamaki, T. & Fukuda, H. *Front. Plant Sci.* **5**, 315 (2014).
5. Milo, R. *et al. Science* **298**, 824–827 (2002).
6. Alon, U. *Nature Rev. Genet.* **8**, 450–461 (2007).
7. Lacayo, C. I. *et al. Plant Physiol.* **154**, 121–133 (2010).

This article was published online on 24 December 2014.

PRECISION MEASUREMENT

Relativity tested with a split electron

Splitting and recombining an electron wave packet has been used to test relativity at a record sensitivity. The result heralds an era of precision measurements of relativity using quantum-information methods. SEE LETTER P.592

V. ALAN KOSTELECKÝ

Lorentz invariance is a fundamental symmetry of space-time that lies at the heart of Albert Einstein's special theory of relativity. Writing in this issue (page 592), Pruttivarasin *et al.*¹ show how they have tested Lorentz invariance for electrons at an unprecedented sensitivity by splitting and recombining a superposition of electron wave functions — an electron wave packet — that is bound to calcium ions. This ingenious experiment opens the door to a new generation of precision tests of relativity using quantum-information techniques.

Lorentz invariance states that the laws of physics that govern a physical system are unchanged for different system orientations

or velocities. Equivalently, the laws of physics exhibit rotation symmetry (spatial isotropy) and boost symmetry. In practice, it is easier to rotate an apparatus than to boost it, and therefore many tests of relativity are designed to explore the spatial isotropy of the behaviour of a physical system. Pruttivarasin and colleagues' experiment can be viewed as a quantum analogue of two famous tests of spatial isotropy: the Michelson–Morley experiment for electrodynamics² and the Hughes–Drever experiment for matter^{3,4}.

The Michelson–Morley experiment uses a device called an interferometer that first splits a light ray into two beams travelling along orthogonal paths (arms), and then reflects and recombines the beams to yield an interference pattern. Monitoring the interference pattern as

the orientation of the apparatus changes can be understood as testing the constancy of the speed of light and hence the rotation symmetry of the laws of electrodynamics. If the speed of light is different in different directions, then the travel times of the two beams vary as the apparatus is rotated, changing the interference pattern.

By contrast, the Hughes–Drever experiment studies the spatial isotropy of the propagation of matter by examining the connection between its energy and its momentum. In essence, the experiment involves placing one or more atoms in a magnetic field, thereby splitting some atomic energy levels and imbuing the system with a definite orientation. Monitoring the frequency of transitions between certain energy levels for a changing orientation of the whole apparatus provides a check of isotropy. If the energy levels depend on the momentum directions of the atomic constituents, the transition frequency will change as the apparatus is rotated.

In their study, Pruttivarasin *et al.* take advantage of the quantum nature of the electron, realizing an electron analogue of the Michelson–Morley and Hughes–Drever experiments through quantum-information techniques that allow a suitable electron wave packet to be created and monitored. The authors confined a pair of calcium ions ($^{40}\text{Ca}^+$), about 16 micrometres apart, in an electromagnetic trap and applied a vertical magnetic field of 3.93 gauss to introduce a definite orientation of the system, which changes as Earth rotates. They then applied laser pulses to electrons bound to the calcium ions, creating an electron wave packet that combines two quantum states of different electron orientations relative to the magnetic field and that oscillates between two configurations. The researchers used further laser pulses to monitor the energy difference between the two states for 23 hours. The experimental procedure can be viewed as repeatedly splitting and then recombining the two quantum states 95 milliseconds later. Violations of spatial isotropy would be seen as variations in this energy difference as Earth rotates.

The experiment confirms spatial isotropy for electrons at the impressive level of one part in 10^{18} . This represents a milestone sensitivity, because it is smaller than the dimensionless ratio of about 10^{-17} between the strengths of the electroweak and gravitational forces that could naturally be expected to govern violations of Lorentz invariance arising in unified theories of quantum physics and gravity⁵. The authors' experiment is thus the first to delve into this realm of sensitivity for electrons.

A crucial subtlety in interpreting tests of Lorentz invariance and hence of special relativity is that a physical reference system is required to define the lengths of rods, the ticking rates of clocks and the idea of orthogonality in space and time. For example, in a classic Michelson–Morley experiment, the lengths

and orthogonality of the interferometer arms are established in terms of properties of matter. The interpretation of the experiment as a test of the isotropy of the speed of light therefore relies on the assumption that these properties are independent of the system's orientation. Analogously, the interpretation of a Hughes–Drever experiment as a test of the rotation symmetry with matter assumes isotropic transition frequencies of light. Indeed, any measurement in physics is really a comparison between two systems, with only the difference between them being physically meaningful.

It follows that Pruttivarasin and co-workers' experiment can be viewed equally as an isotropy test for light assuming conventional electrons or as an isotropy test for electrons assuming conventional electrodynamics. In the first scenario, the experiment is interpreted as searching for possible spatial anisotropies in the Coulomb force that binds the electron wave packet to the calcium ions, and hence in the laws of electrodynamics, and the results represent a fivefold improvement in sensitivity over current limits^{6,7}. In the second picture, the energy of the electron wave packet depends on the direction of its momentum, and the new constraints sharpen existing bounds⁸ 100-fold.

Possible experimental improvements include choosing different ions to yield a longer-lived electron wave packet, binding the wave packet to ions of greater charge, preparing the wave packet more directly, and taking data over a longer period. Another 100-fold improvement in sensitivity may lie within reach.

Other tests of Lorentz invariance are also feasible using these methods. Lorentz

violations accessible in principle include those characterized by nine coefficients⁹, each corresponding to a different physical effect. Six coefficients govern violations of rotation symmetry, whereas three control boost violations. Pruttivarasin *et al.* obtained constraints involving four of the six types of rotation-symmetry violation (see Table 1 of the paper¹). The remaining two could be studied in a similar experiment mounted on a turntable with its axis of rotation differing from that of Earth. The three coefficients controlling boost violations could also be measured with tenfold improved sensitivity using data acquired over many months, by taking advantage of the changing direction of Earth's velocity as it revolves around the Sun. Stay tuned for future cutting-edge tests of relativity using these quantum-information techniques. ■

V. Alan Kostelecký is in the Physics Department and the Center for Spacetime Symmetries, Indiana University, Bloomington, Indiana 47405, USA.
e-mail: kostelec@indiana.edu

1. Pruttivarasin, T. *et al.* *Nature* **517**, 592–595 (2015).
2. Michelson, A. A. & Morley, E. W. *Am. J. Sci.* **34**, 333–345 (1887).
3. Hughes, V. W., Robinson, H. G. & Beltran-Lopez, V. *Phys. Rev. Lett.* **4**, 342–344 (1960).
4. Drever, R. W. P. *Phil. Mag.* **6**, 683–687 (1961).
5. Kostelecký, V. A. & Potting, R. *Phys. Rev. D* **51**, 3923–3935 (1995).
6. Herrmann, S. *et al.* *Phys. Rev. D* **80**, 105011 (2009).
7. Eisele, Ch., Nevsky, A. Yu. & Schiller, S. *Phys. Rev. Lett.* **103**, 090401 (2009).
8. Hohensee, M. A. *et al.* *Phys. Rev. Lett.* **111**, 050401 (2013).
9. Colladay, D. & Kostelecký, V. A. *Phys. Rev. D* **58**, 116002 (1998).

GENOMICS

CRISPR engineering turns on genes

The repurposing of a bacterial defence system known as CRISPR into a potent activator of gene expression in human cells enables powerful studies of gene function, as exemplified in cancer cells. [SEE ARTICLE P. 583](#)

SEUNG WOO CHO & HOWARD Y. CHANG

The ability to turn on any gene at will has been a long-held dream of molecular biologists. Most genes are dynamically turned on and off by specific biological processes, and manipulation of the level of gene expression is a key method for studying the functions of each gene, regulatory element and pathway. On page 583 of this issue, Konermann *et al.*¹ describe an elegant strategy for converting the CRISPR/Cas9 system — a bacterial defence system against foreign DNA —

into a potent and selective gene activator. The authors demonstrate how this approach can be used to test the effects of turning on tens of thousands of individual genes in parallel.

CRISPR, which stands for clustered regularly interspaced short palindromic repeats, is the name given to regions of bacterial DNA encoding RNA sequences that recognize foreign DNA sequences, such as those in viruses, through direct sequence complementarity. In bacteria, these guide RNAs assemble into complexes with the enzyme Cas9 or other proteins that specifically cut the recognized

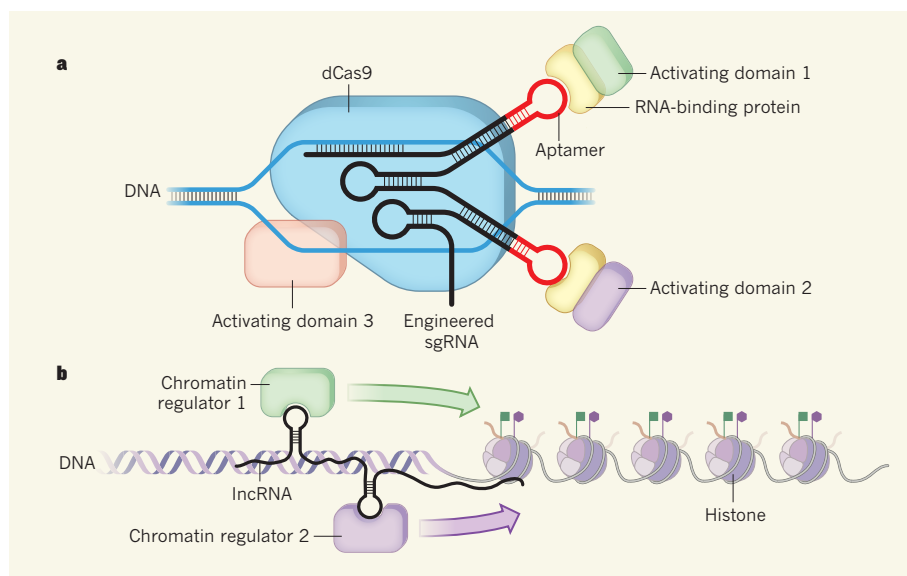


Figure 1 | RNAs as modular scaffolds for gene regulation. **a**, Konermann *et al.*¹ have engineered natural CRISPR short guide RNA molecules (sgRNAs) to become modular scaffolds for protein assembly, thus forming complexes that can specifically regulate gene expression through the action of multiple transcription-regulatory domains. Aptamer structures in the sgRNAs recruit specific proteins: in this case, RNA-binding proteins fused to the transcription-activating domains of different transcription factors. The complex's DNA-targeting specificity is provided by base pairing between a different part of the CRISPR RNA and the target DNA sequence. The enzyme dCas9, which associates with CRISPR RNA and helps to unwind the DNA, can also be fused to another regulatory domain, thereby increasing the diversity of effects. **b**, This engineering of the CRISPR system incorporates design principles from natural long non-coding RNAs (lncRNAs), which can recruit multiple cellular machines that regulate gene expression by modifying chromatin (the complex of histone proteins and DNA in the cell nucleus).

DNA sequence, thereby destroying it and protecting the bacterium from invasion². CRISPR/Cas9 thus represents a programmable DNA-targeting system, with its specificity determined by the RNA sequence.

Molecular biologists have adopted this system to allow the rapid mutation or replacement of genomic sequences — a strategy called genome editing. A breakthrough came in 2012, when the system was simplified to use single short guide RNA (sgRNA) molecules to programme CRISPR specificity³. Soon after, it was found that a mutant of Cas9 that no longer cuts DNA, termed dCas9, can be used as a DNA-binding platform^{4,5}. When dCas9 is linked to portions (domains) of proteins involved in transcriptional activation or repression and then targeted, using CRISPR, to promoter sequences that regulate transcription of particular genes, these fusion proteins can modulate natural gene-expression levels^{4,5}. However, the change in gene expression achieved by this approach is too low — less than or around fivefold activation — for many applications.

Konermann and colleagues overcame this low efficiency of gene activation by turning the CRISPR sgRNA into a modular platform for assembling multiple different transcriptional activators (Fig. 1a). They identified two regions of the sgRNA that can be appended with short sequences that attract an RNA-binding protein, which is in turn fused to the transcription-activation domains of different mammalian transcription factors. The

authors termed this system the synergistic activation mediator (SAM), and demonstrate that it induced more than 100-fold activation of 12 genes that were not efficiently activated by the dCas9-activator fusion protein.

To illustrate the potential applications of this approach, Konermann *et al.* created a library of engineered sgRNAs that allowed more than 23,000 human genes to be individually turned on. They then asked which genes, on activation, give melanoma cancer cells the ability to escape the killing effects of the drug PLX-4720, a mainstay of melanoma treatment. The degree of drug resistance conferred by turning on different genes was determined by the relative frequency of sgRNAs in the melanoma cells after drug treatment. The highly enriched sgRNAs included those corresponding to genes involved in known drug-resistance pathways and to genes that are expressed at increased levels in patients with drug-resistant melanoma — verifying that the SAM method can identify biologically relevant outcomes of altered gene expression.

The success of this CRISPR engineering effort has direct parallels with natural mechanisms of gene regulation. Enhancers are DNA sequences that turn on gene expression, and they typically contain recognition sequences for several different types of transcription factor. Moreover, enhancers and other regulatory elements often generate long non-coding RNAs (lncRNAs), which act as modular scaffolds to recruit diverse cellular machines that

modify chromatin (the complex of histone proteins and DNA in the cell nucleus) and thereby regulate gene expression (Fig. 1b). lncRNAs are molecular recipes writ large, containing both instructions for the set of biochemical activities to be assembled and the genomic address at which these activities should be carried out⁶.

Two other recent studies have also demonstrated improved efficiency or flexibility of CRISPR-guided gene regulation, through repurposing dCas9 or CRISPR RNA as modular scaffolds^{7,8}. One of those studies⁸ further showed that multiple engineered CRISPR RNAs can simultaneously turn different genes in the same cell on and off to manipulate a metabolic pathway. Together with Konermann and colleagues' findings, these studies show that mimicking natural lncRNAs is an efficient way to orchestrate multiple proteins to work together across the genome. It may be possible to deliver defined combinations of various effector proteins to the same genomic location using one sgRNA molecule. Future construction of novel multifunctional artificial proteins or non-coding RNAs will also be worthwhile, given the broad usefulness of such tools in biotechnology.

This next generation of CRISPR technology opens the door to studying the functions of many genes and DNA sequences. RNA-interference techniques have been widely used for studying the effects of loss of gene function over the past decade, but this approach can yield a high rate of false-positive results due to nonspecific targeting. Meanwhile, gain-of-function studies using overexpression techniques may not recapitulate normal RNA regulatory processes, such as alternative splicing. Artificial transcription factors based on DNA-binding proteins called zinc fingers and TALENs are alternatives for altering gene expression, but these are difficult to construct on a genome-wide scale. Thus, the comprehensive coverage of CRISPR libraries and the modular nature of this approach are strong advantages over other techniques. However, CRISPR targeting may also have off-target effects⁹, and additional validation experiments may be needed to confirm any effects of altered gene expression identified using this approach.

In their melanoma-cell experiments, Konermann *et al.* identified 13 genes whose altered expression was individually sufficient to confer drug resistance. However, diseases or profound biological effects often result from complex regulation of multiple genes at the same time — a good example is the finding that four genes must be expressed together for the generation of induced pluripotent stem cells (a form of stem cell generated from adult cells)¹⁰. Thus, we will need a detailed understanding of regulatory networks and will need to experiment with gene sub-libraries and dosages to identify the sets of genes that

together determine certain characteristics. The CRISPR/Cas system will be a versatile tool for this purpose, owing to its capacity for multiplexed targeting and, now, multiplexed deployment of diverse effector domains. ■

Seung Woo Cho and Howard Y. Chang are in the Program in Epithelial Biology and the

Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305, USA.
e-mail: howchang@stanford.edu

1. Konermann, S. *et al. Nature* **517**, 583–588 (2015).
2. Doudna, J. A. & Charpentier, E. *Science* **346**, 1258096 (2014).
3. Jinek, M. *et al. Science* **337**, 816–821 (2012).

4. Gilbert, L. A. *et al. Cell* **154**, 442–451 (2013).
5. Mali, P. *et al. Nature Biotechnol.* **31**, 833–838 (2013).
6. Tsai, M. C. *et al. Science* **329**, 689–693 (2010).
7. Gilbert, L. A. *et al. Cell* **159**, 647–661 (2014).
8. Zalatan, J. G. *et al. Cell* **160**, 339–350 (2014).
9. Wu, X. *et al. Nature Biotechnol.* **32**, 670–676 (2014).
10. Takahashi, K. & Yamanaka, S. *Cell* **126**, 663–676 (2006).

EARTH SCIENCE

Free and forced climate variations

A combination of simulations and data shows that short-term climate trends are dominated by natural internal variations, providing a basis for climate forecasting, but not for assessing sensitivity to forced changes. SEE ARTICLE P.565

JAMES RISBEY

The global mean surface air temperature of the planet has increased over the past century, punctuated by periods of faster and slower warming lasting a decade or more at a time¹ (Fig. 1). The apparent 'slowdown' in the rate of surface warming in the past 15 years is common in the record (as are accelerations), but has nonetheless drawn questions about the sources of fluctuations in the warming rate and about the implications of these fluctuations for climate-model projections. On page 565 of this issue, Marotzke and Forster² describe how they have used observations and climate-model runs to show that these fluctuations are dominated by natural (free) variations, and that decadal-scale trends do not provide grounds for revising the sensitivity of climate models to forcings such as increased concentrations of greenhouse gases.

It has long been known that the global surface temperature fluctuates on multidecadal time-scales in response to natural variations in ocean circulation and other processes. Since at least the 1930s, climatologists have used 30-year averages as a standard climate 'normal' to smooth out decadal variations³. The 1995 Intergovernmental Panel on Climate Change report⁴ identified volcanic activity, variations in incoming solar radiation, and ocean circulation as drivers of decadal and longer variability in global surface temperature. An innovation of Marotzke and Forster's work is their use of coupled ocean–atmosphere

climate-model runs⁵ to extract the contributions of processes internal to the climate system (free variation) that are due mainly to ocean circulation and its coupling with the atmosphere, and the contributions of external processes (forced variation) such as volcanoes, variations in greenhouse-gas concentrations and solar output.

Marotzke and Forster provide the first quantification of these contributions for trends of length 15 years and 62 years in the century-long instrumental climate record. Their method uses the statistical tool of multiple regression on components of the surface energy balance (relating changes in forcing and temperature) to fit global surface-temperature trends in models, on the basis of the model forcing and response. The method

does not account for any systematic errors in climate-model representations of forcing, and may under-represent some slow climate responses that alter temperatures on time-scales longer than the trend periods examined. Furthermore, the method attributes all residual between the model-fitted trends and the actual model trends to free variation. These assumptions mean that the method may over-estimate the free contribution to the model trends, but trends predicted in this way are consistent with those observed. The results show that the free variations are the dominant contributor to model 15-year trends, and are also important, but less so than forced variations, for the model 62-year trends. The results provide model-based confirmation that free variations dominate 15-year trends across the instrumental period.

The other big issue addressed by Marotzke and Forster is whether the coupled climate models have overestimated the sensitivity of the climate system to forced changes. Some have claimed that the recent slowdown in surface warming is not well captured in climate-model projections⁶, and further, that it implies that climate models may have over-estimated the response of the climate system to increases in greenhouse gases^{1,7}.

A range of studies have now addressed the first claim using a variety of approaches. Some of these address the issue that a single 15-year period in the real world is not by design synchronized with the same 15 years in model projections. The free variations in the latter are not, and cannot be, synchronized with the real world⁸. This means that neither a single model run, nor the average of many model runs, is expected to match a given observed 15-year trend. However, by selecting model runs according to criteria that effectively phase-lock the model free variation to the observed free variation, the models provide good representations of observed 15-year trends^{8,9}.

The observed temperature trends in the most recent 15-year period in which the slowdown occurs can also be reproduced in the models by specifying the observed surface winds to

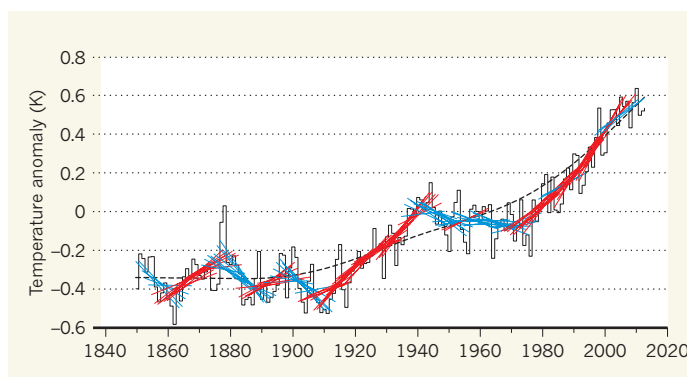


Figure 1 | Fifteen-year temperature trends. The solid black line is the time series of the global mean surface temperature¹⁴, plotted as a departure (anomaly) from a baseline period 1961–90. The dashed black line is a smooth fit to this series, representing the long-term warming rate. The blue and red lines are linear trends for each 15-year segment running over 1850–64, 1851–65, ..., 1999–2013. Each 15-year segment is shown in red if the trend rises faster than the long-term warming rate in the same 15-year period and blue if it rises more slowly. Marotzke and Forster² show that these 15-year trends are dominated by natural (free) variations. The free variations drive the 15-year trends above and below the long-term warming rate as they ride along with it.

synchronize the model free variation¹⁰. Other work shows that, by examining all 15-year trends in observations in the instrumental record and a large ensemble of 15-year trends in climate projections, there is no systematic bias in the model trends^{2,8}.

The response of a climate model and the climate system over a set period depends on the length of the period examined. For longer periods, the change in forcing produces a change in temperature, which is in turn amplified through other responses of the climate system, primarily related to the phases of water (vapour, cloud, ice). The amplification process is termed the climate feedback. If climate feedbacks are too strong in a model, then the model may overestimate the forced response over a period.

Using measured feedbacks and forcings in climate models¹¹, Marotzke and Forster show that the contribution of spread in climate feedback to spread in model temperature trends over 15-year periods is small compared with those from free variations and direct forcing. This means that comparisons of observed and model 15-year trends can yield information about free variations in the model, but not about climate feedback or sensitivity, because climate feedback plays little part on this timescale. Even allowing for some uncertainty in Marotzke and Forster's method, the results are emphatic enough for there to be little room for concluding that climate models are overestimating the response of the climate system

on the basis of short-trend comparisons. Such ad hoc revisions of climate sensitivity are also problematic because changes in forcing associated with aerosol emissions are hard to quantify (confounding the forced change), and accurate estimates of ocean heat content are not available before about 2006.

With the work of Marotzke and Forster, we now have a clearer view of the contribution of free and forced variations to 15-year trends. The climate system and climate models generate a long-term warming in response to steadily growing forcing by greenhouse-gas increases. On shorter timescales, the warming is not steady, but speeds up and slows down (Fig. 1) in response to free variations in the climate system, and much less so due to forced variations from solar cycles, volcanic eruptions and other aerosol sources. The (non-greenhouse-gas) forced variations can have a larger role in any particular 15-year period, but because they are irregular (volcanoes) and/or weak (solar variations), their role is generally much smaller when viewed across many 15-year periods.

The source of the free variations is in the intrinsic circulation modes in the ocean and their coupling to intrinsic modes in the atmosphere¹². Some of this variation is predictable and is stimulating the development of initial-ized climate forecasts¹³. These hold out the promise of better adaptation to climate variability on short timescales. The free variations and occasional enhanced forcing by aerosols

and solar cycles drive fluctuations that ride on top of the longer-term warming response but do not subdue it. ■

James Risbey is in the Oceans and Atmosphere Flagship, Commonwealth Scientific and Industrial Research Organisation, Hobart, Tasmania 7001, Australia.

e-mail: james.risbey@csiro.au

1. Stocker, T. F. et al. in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. F. et al.) 33–115 (Cambridge Univ. Press, 2013).
2. Marotzke, J. & Forster, P. M. *Nature* **517**, 565–570 (2015).
3. Arguez, A. & Vose, R. S. *Bull. Am. Met. Soc.* **92**, 699–704 (2011).
4. Houghton, J. T. et al. (eds) *Climate Change 1995: The Science of Climate Change* (Cambridge Univ. Press, 1996).
5. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. *Bull. Am. Met. Soc.* **93**, 485–498 (2012).
6. Fyfe, J. C., Gillett, N. P. & Zwiers, F. W. *Nature Clim. Change* **3**, 767–769 (2013).
7. Otto, A. et al. *Nature Geosci.* **6**, 415–416 (2013).
8. Risbey, J. S. et al. *Nature Clim. Change* **4**, 835–840 (2014).
9. Meehl, G. A., Teng, H. & Arblaster, J. M. *Nature Clim. Change* **4**, 898–902 (2014).
10. England, M. H. et al. *Nature Clim. Change* **4**, 222–227 (2014).
11. Forster, P. M. et al. *J. Geophys. Res.* **118**, 1139–1150 (2013).
12. O'Kane, T. J. et al. *Ocean Model.* **69**, 1–21 (2013).
13. Meehl, G. A. et al. *Bull. Am. Met. Soc.* **95**, 243–267 (2014).
14. Cowtan, K. & Way, R. G. *Q. J. R. Met. Soc.* **140**, 1935–1944 (2014).

CANCER

Risk factors and random chances

The discovery that the estimated number of stem-cell divisions in a tissue correlates with cancer incidence suggests that the varying probability of developing cancer in different tissues is mostly down to random mutations.

DOMINIK WODARZ & ANN G. ZAUBER

Cancer arises through the accumulation of molecular changes that together allow cells to grow in an uncontrolled manner. Many factors contribute to this process, including hereditary genetic mutations and environmental hazards, such as exposure to smoking or radiation. But cancer can still emerge in the absence of these factors, as a result of random mutations that arise during cell division¹. Furthermore, little is known about how each risk factor affects different tissues, in which cancers arise at varying frequencies. Writing in *Science*, Tomasetti and Vogelstein² show that about 65% of the

differences in cancer incidence between tissues can be simply explained by the estimated total number of stem-cell divisions in those tissues. This result suggests that, rather than environmental and heredity influences, stochastic accumulation of mutations during DNA replication is the major cause of variations in cancer incidence between tissues.

Estimating the number of stem-cell divisions that occur in a tissue is a complex task. Tomasetti and Vogelstein performed an extensive literature search to obtain information about stem-cell numbers in various tissues. They then used mathematical and statistical models to estimate the number of stem-cell divisions in each tissue during the

average human lifespan, and correlated this with publicly available cancer-incidence data. Although some cancers are certainly driven by genetic predisposition or environmental risk factors, the authors' analysis indicated that these factors explain only one-third of the variation in cancer incidence between tissues (Fig. 1). The rest of the variability was explained by the different number of stem-cell divisions estimated to occur in different tissues. However, some commonly occurring cancers, such as breast and prostate, were not included in the analysis, and these must be examined when data become available. Importantly, although this paper attempts to explain the variation in cancer incidence among tissues, it does not try to quantify the percentage of cancers that arise owing to random accumulation of mutations alone, which is a different measure.

This study suggests that processes governing the evolution of cancer cells need to be better understood, so that they can be manipulated to delay the onset of cancer. One drug that might affect evolutionary dynamics is aspirin, which protects against a variety of cancers³. Aspirin modulates evolutionary parameters that determine the rate of cancer development, including cell-division and cell-death rates⁴, and the persistence of cells with elevated mutation rates⁵.

In addition to experimental studies, math-

ematical models are crucial for investigating the evolutionary dynamics of cancer⁶. Such models have led to the development of several principles that define how random mutations and subsequent selection influence the emergence and growth of tumour cells in certain settings. For example, stochastic evolutionary models have been used to study the emergence and spread of tumour cells in healthy tissue^{7,8}. Furthermore, models that take into account stem-cell dynamics have been used to study the evolutionary pathways by which cancerous cells escape the constraints of tissue regulation⁹.

Nonetheless, much remains to be done, and Tomasetti and Vogelstein's paper highlights some intriguing directions for future research. For instance, although the authors assumed that random cancer-causing mutations occur during stem-cell division, it is actually unknown in which tissues stem cells (rather than more-differentiated cells) become cancerous¹⁰. The correlation they report could arise either way, because the number of stem-cell divisions probably correlates with the number of divisions undertaken by the differentiating daughters of stem cells.

Another gap in our knowledge concerns the relative importance for cancer evolution of embryonic cell divisions compared with divisions after birth¹¹. Although many cell divisions occur after birth in most tissues included in this study, almost no cell division occurs after birth in the cells that give rise to glioblastoma¹², an aggressive brain tumour. But Tomasetti and Vogelstein found that the incidence of glioblastoma fits the same trend as the other cancers they studied.

Last but not least, we need to understand how various environmental selection pressures influence the fate of the mutant cells that are generated by chance. Microenvironmental conditions, such as immune responses, inflammation or the presence of cancer-causing molecules, can affect the fitness of specific mutants and thus their ability to grow and give rise to disease. These conditions can change during ageing, resulting in environments that are conducive to the growth of cancerous cells¹³. Although the correlation observed by Tomasetti and Vogelstein is certainly intriguing, and although accumulation of mutations clearly has a central role in causing variability between tissues, much research is needed to disentangle the complex, multifactorial interactions that result in disease.

A landmark paper¹⁴ published in 1981 suggested that most cancers could be averted by removing various lifestyle, behavioural and environmental risk factors prevalent in the population, and thus that much of the risk of cancer could be controlled. Risk factors are certainly involved in promoting the occurrence of many cancers, especially the more common ones. It therefore makes sense to use the presence or absence of risk factors to identify individuals who should

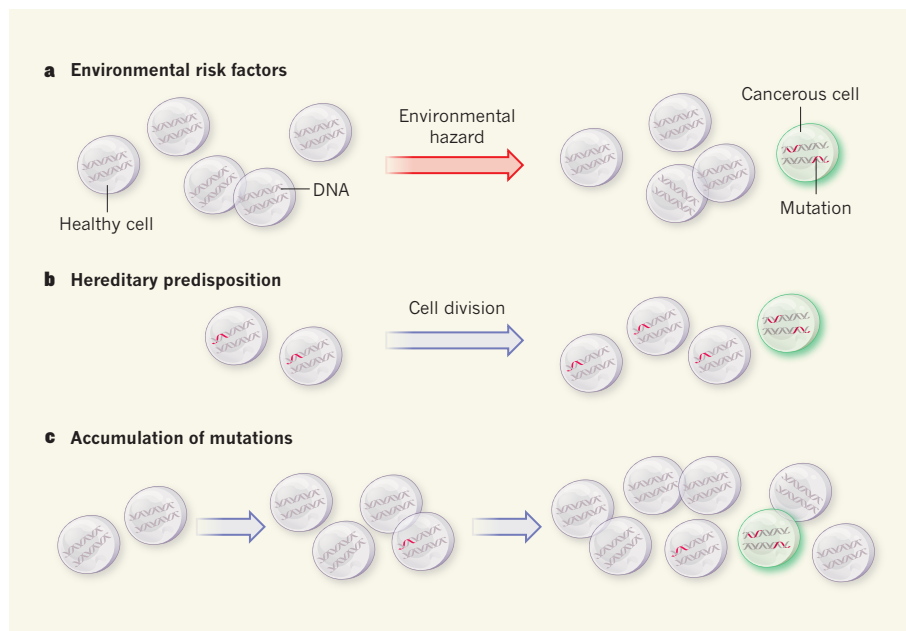


Figure 1 | Mechanisms of cancer development. **a**, Exposure to environmental carcinogens can cause genetic mutations in a healthy cell population, leading to the generation of cancerous cells (two mutations are shown as sufficient for tumour growth here, although more are often required). **b**, If a mutation associated with cancer has been passed down from parents to offspring, the offspring has a hereditary predisposition to that cancer. Tumour growth is initiated if any of these cells randomly acquires a second mutation during cell division. **c**, In the absence of environmental risk factors and genetic predisposition, acquisition of random mutations alone can be sufficient to cause cancer. Tomasetti and Vogelstein² report that this is the major cause of the variable rates at which cancer arises in different tissues.

be screened for cancer. If, however, random genetic changes are more-relevant drivers of carcinogenesis, then biomarkers and early-detection methods will have to be developed to prevent cancer mortality in the general population.

This might be problematic. Screening for many of the cancers included in Tomasetti and Vogelstein's study is currently difficult, particularly for rare cancers. Screening is successful only when several criteria are fulfilled: that tests are available that detect early disease; that the test's sensitivity and specificity for the given cancer is high; that people are willing to be screened; and that effective treatments exist that can be applied to early-stage cancers to prevent death. Furthermore, the benefits of screening must outweigh the risks of offering screening to the general population. The 'number needed to screen' to prevent one cancer death must be feasible, given the resources required to perform screens on the population as a whole (see go.nature.com/61efy9).

In summary, Tomasetti and Vogelstein's findings emphasize the role of basic evolutionary mechanisms in cancer development, and might lead to new chemoprevention therapies that slow the evolutionary processes at work. Although screening according to risk factors remains a crucial intervention strategy, an improvement in our cancer-prevention efforts in the general population will require the generation of early-detection techniques. ■

Dominik Wodarz is in the Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697, USA.

Ann G. Zauberman is in the Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA.

e-mails: dwodarz@uci.edu; zauberama@mskcc.org

1. Vogelstein, B. & Kinzler, K. W. (eds) *The Genetic Basis of Human Cancer* (McGraw-Hill, 2002).
2. Tomasetti, C. & Vogelstein, B. *Science* **347**, 78–81 (2015).
3. Rothwell, P. M. et al. *Lancet* **377**, 31–41 (2011).
4. Goel, A., Chang, D. K., Ricciardiello, L., Gasche, C. & Boland, C. R. *Clin. Cancer Res.* **9**, 383–390 (2003).
5. Rüschhoff, J. et al. *Proc. Natl Acad. Sci. USA* **95**, 11301–11306 (1998).
6. Wodarz, D. & Komarova, N. L. *Dynamics of Cancer: Mathematical Foundations of Oncology* (World Scientific, 2014).
7. Nowak, M. A. et al. *Proc. Natl Acad. Sci. USA* **99**, 16226–16231 (2002).
8. Nowak, M. A., Michor, F. & Iwasa, Y. *Proc. Natl Acad. Sci. USA* **100**, 14966–14969 (2003).
9. Rodriguez-Brenes, I. A., Komarova, N. L. & Wodarz, D. *Proc. Natl Acad. Sci. USA* **108**, 18983–18988 (2011).
10. Komarova, N. L. & Wang, L. *Cell Cycle* **3**, 1558–1565 (2004).
11. Frank, S. A. & Nowak, M. A. *Nature* **422**, 494 (2003).
12. Spalding, K. L., Bhardwaj, R. D., Buchholz, B. A., Druid, H. & Frisén, J. *Cell* **122**, 133–143 (2005).
13. Rozhok, A. I., Salstrom, J. L. & DeGregori, J. *J. Aging* **6**, 1033–1048 (2014).
14. Doll, R. & Peto, R. *J. Natl Cancer Inst.* **66**, 1192–1308 (1981).

Forcing, feedback and internal variability in global temperature trends

Jochem Marotzke¹ & Piers M. Forster²

Most present-generation climate models simulate an increase in global-mean surface temperature (GMST) since 1998, whereas observations suggest a warming hiatus. It is unclear to what extent this mismatch is caused by incorrect model forcing, by incorrect model response to forcing or by random factors. Here we analyse simulations and observations of GMST from 1900 to 2012, and show that the distribution of simulated 15-year trends shows no systematic bias against the observations. Using a multiple regression approach that is physically motivated by surface energy balance, we isolate the impact of radiative forcing, climate feedback and ocean heat uptake on GMST—with the regression residual interpreted as internal variability—and assess all possible 15- and 62-year trends. The differences between simulated and observed trends are dominated by random internal variability over the shorter timescale and by variations in the radiative forcings used to drive models over the longer timescale. For either trend length, spread in simulated climate feedback leaves no traceable imprint on GMST trends or, consequently, on the difference between simulations and observations. The claim that climate models systematically overestimate the response to radiative forcing from increasing greenhouse gas concentrations therefore seems to be unfounded.

The GMST has risen in the past fifteen years at a rate that is only one-third to one-half of the average over the second half of the twentieth century (see, for example, refs 1–5). This hiatus is not reproduced in most simulations with present-generation climate models, which instead over the period 1998–2012 show a larger GMST trend than observed^{5–14}. The difference between GMST observations and simulations is caused in part by quasi-random internal climate variability^{5–10,13,14}, which arises because of chaotic processes in the climate system. But part of the difference is probably caused by errors in the model radiative forcing^{5,12,14–16} or in the model response to radiative forcing^{5,14,17,18}. The relative magnitudes of these three contributions are poorly known. Here we quantify how forcing, feedback and internal climate variability contribute to spread in simulated historical GMST trends and, hence, to the differences between models and observations.

We use a three-pronged approach. First, we note that, owing to quasi-random internal climate variability, the difference between observed and simulated trends likewise contains quasi-random contributions. To avoid focusing too strongly on the particular period 1998–2012, which contains some climate extremes relevant for GMST^{19–21} and is hence unlikely to be reproduced in a simulation containing quasi-random contributions, we analyse GMST trends of a certain length for the entire period 1900–2012¹³. Second, we quantify the contributions of forcing, climate feedback, ocean heat uptake and internal variability to simulated GMST trends, through a multiple linear regression approach that is physically motivated by the global surface energy balance. And, third, we investigate trends over both 15 and 62 years, representing decadal and multidecadal timescales, respectively. We combine these three aspects into a new unified conceptual framework, which allows us to put the GMST trends over the 15-year period 1998–2012 into the appropriate context.

We first create linear trends from an ordinary least-squares fit, and perform all statistical analyses on these trends. This procedure implies that the analysis must be repeated for each trend length, in contrast to previous work aiming at attributing elements in the observed GMST time series itself; such elements include effects of volcanic eruptions,

solar variability, anthropogenic forcing, El Niño events and sources of atmospheric dynamic variability including land–sea contrasts^{13,14,22–25}. Because the amplitude of internal variability decreases with increasing trend length^{3,26}, we expect a clearer breakdown into the individual contributions from forcing, feedback and internal variability if we focus on one trend length at a time. We analyse trends over both 15 and 62 years, because these were the trend lengths primarily considered in the Intergovernmental Panel on Climate Change Assessment Report 5⁵ (AR5).

Observed and simulated 15-year trends

To gauge whether the difference between simulations and observations is unusual over the hiatus period, we first compare observed and simulated 15-year trends over the entire period 1900–2012 (Fig. 1; see also ref. 13). We use the HadCRUT4 observational data set²⁷ and the ‘historical’ simulations conducted under the auspices of the Coupled Model Intercomparison Project Phase 5²⁸ (CMIP5), extended for the years 2006–2012 with the RCP4.5 scenario runs (Extended Data Fig. 1 and Extended Data Table 1). The simulation output is subsampled using the HadCRUT4 data mask¹¹, to account for the effects of incomplete observational coverage^{29,30}.

Figure 1a contains the joint relative frequency distribution of 15-year GMST trends across the 114 available CMIP5 simulations, as a function of start years since 1900 and trend size. Compared with the CMIP5 ensemble, observed trends are distributed in no discernibly preferred way and occur sometimes at the upper end of the ensemble (for example, for start year 1927 the best-estimate observed trend is larger than 110 of the 114 simulated trends; Fig. 1b) and sometimes at the lower end of the ensemble (for example, for start year 1998 the best-estimate observed trend is smaller than all 114 simulated trends; Fig. 1c)^{5,13,26}.

In both cases depicted in Figs 1b, c, fewer than 5% of the simulations lie in one of the tails relative to the observed trend. Hence, if a 5% criterion for statistical significance were used, one would diagnose formal model–observation inconsistency for 15-year trends with start years 1927 and 1998¹¹. But when the comparison is repeated for all start years, the rank that the observed trend would have as a member of the

¹Max Planck Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg, Germany. ²School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK.

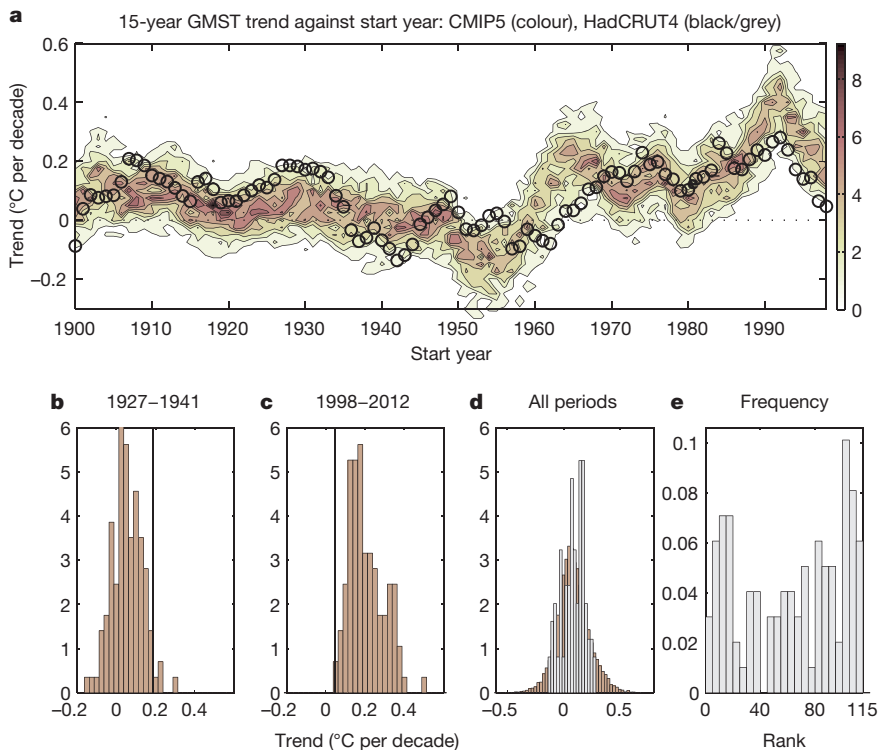


Figure 1 | Simulated and observed 15-year GMST trends since 1900. **a**, Joint relative frequency distribution of GMST trends as a function of start year and trend size, based on the full 114-member ensemble (in bins of 0.025 °C per decade, as shown by colour scale). Circles mark the observed trend from the HadCRUT4 data set²⁷. **b**, Vertical cross-section of **a** for start year 1927; vertical line marks the observed trend. **c**, As **b**, but for start year 1998. **d**, Marginal distribution of simulated GMST trend as a function of trend size (purple), obtained by time-averaging the joint distribution in **a**, and observed trend distribution (grey). **e**, Frequency distribution of the rank that the observed trend would have as a member of the model ensemble (rank 1, observed trend smaller than all simulations; rank 115, observed trend larger than all simulations; bin size is five. All histograms are normalized such that their area integral is unity. In **a**, each vertical cross-section is normalized.

ensemble of simulated trends³¹ shows no apparent bias (Fig. 1e), indicating that the observed and simulated distributions of 15-year trends are broadly consistent with each other. Any position of the observed trend within the ensemble of simulated trends—including a position at or near the margin—is thus dominated by quasi-random effects (although for any particular start year, a non-negligible contribution from systematic errors cannot be excluded).

The marginal distribution of simulated GMST trends as a function of trend size is wider than the observed distribution of trends (Fig. 1d), a finding consistent with that from the previous generation of climate models³². The width is exaggerated owing to contributions arising at three distinct periods. Some simulated trends with start years from around 1950–1960 are more strongly negative than any observed trends since 1900, and some simulated trends with start years from around 1960–1970 and from around 1985–1998 are more strongly positive than any observed trends since 1900 (Fig. 1a). All three periods (1950–1960, 1960–1970 and 1985–1998) are influenced by volcanic eruptions (Mount Agung in 1963 and Mount Pinatubo in 1991). We speculate that some, though not all, models overestimate the cooling induced by an eruption and the subsequent warming recovery (see, for example, ref. 12 concerning a confounding role of El Niño).

The mean over all simulated 15-year trends during the period 1900–2012 is 0.086 ± 0.001 °C per decade (mean \pm s.e.m.; $n = 11,186$), in excellent agreement with the observed 0.088 ± 0.01 °C per decade ($n = 99$). Furthermore, of all 11,186 pairwise comparisons that are possible between simulated and observed trends, the observed trend is higher in 53.6% of cases, which is slightly above the 50% expected for a perfectly unbiased model ensemble. Figure 1 demonstrates that when viewed over the entire period 1900–2012, the 15-year GMST trends simulated by the CMIP5 ensemble show no systematic deviation from the observations.

Our interpretation of Fig. 1 tacitly assumes that the simulated multimodel-ensemble spread accurately characterizes internal variability, an assumption shared with other interpretations of the position of observed trends relative to simulated trends (for example the reduction in Arctic summer sea ice^{5,33,34}). We now test the validity of this assumption, by identifying deterministic and quasi-random causes of ensemble spread. We exploit the availability of a large number of simulations—114 realizations with 36 different models, with forcing information

available for 75 realizations with 18 different models³⁵ (Extended Data Figs 1 and 2 and Extended Data Table 1)—and investigate the contributions of radiative forcing, climate feedback and ocean heat uptake to all simulated 15-year and 62-year GMST trends during the period 1900–2012.

Energy balance and multiple regression

Our starting point is the globally averaged energy balance for the surface layer^{35–37}. An increasing trend ΔF in effective radiative forcing (ERF) causes an increasing trend ΔT in GMST. This in turn leads to increased outgoing radiation, which in linearized form is written as $\alpha \Delta T$, where α is the climate feedback parameter. Furthermore, the GMST increase leads to increased heat transfer from the surface layer to the subsurface ocean, written, again in linearized form, as $\kappa \Delta T$, where κ is the ocean heat uptake efficiency. The thermal adjustment of the surface layer to ΔF is expected to occur within a few years^{35–37}. This means that for timescales of one to several decades, the surface energy balance is in quasi-steady state and reads

$$(\alpha + \kappa) \Delta T = \Delta F$$

which produces the energy-balance ‘prediction’ for the GMST trend:

$$\Delta T = \Delta F / (\alpha + \kappa) \quad (1)$$

Each CMIP5 model simulates its own ERF time series over the historical period. These time series were diagnosed previously³⁵; if multiple realizations were available for a model, the ensemble average of the individual diagnosed ERF time series for this model was given³⁵ and is used here. The individual α and κ values were previously determined for each CMIP5 model from a regression of global top-of-atmosphere energy imbalance against GMST^{5,35,38–41}, in turn based on simulations in which the CO₂ concentration was quadrupled abruptly. The ranges of α and κ are 0.6–1.8 and 0.45–1.52 W m^{–2} °C^{–1}, respectively. That α and κ in the CMIP5 models might vary with time and climate state^{42,43} is ignored here. There is some positive, though not statistically significant, correlation between α and κ (across the 75-member subensemble, the correlation is 0.17 with $P = 0.14$).

Each model's α value is related to its equilibrium climate sensitivity (ECS) by

$$\text{ECS} = F_{2x}/\alpha \quad (2)$$

where F_{2x} is the effective radiative forcing from a doubling of the pre-industrial atmospheric CO_2 concentration. The reference value for F_{2x} is 3.7 W m^{-2} (see, for example, ref. 44), but F_{2x} varies between 2.6 and 4.3 W m^{-2} across the CMIP5 ensemble^{5,38}. To avoid confounding the uncertainty in model response with the uncertainty from CO_2 forcing, we use α and not ECS to characterize model response.

On the basis of the physical foundation of energy balance (equation (1)), we determine the extent to which the across-ensemble variations of ΔF , α and κ contribute to the ensemble spread of GMST trends ΔT , using the 75-member subensemble of CMIP5 historical simulations for which radiative forcing information can be obtained from the CMIP5 archive³⁵ (Extended Data Table 1). The presence of internal variability is included in our framework by adding a random term to equation (1), such that our equation is

$$\Delta T = \Delta F/(\alpha + \kappa) + \varepsilon \quad (3)$$

Because equation (3) assumes an increasing trend in ERF, its validity is somewhat questionable following a volcanic eruption (see, for example, ref. 25). However, Extended Data Fig. 3 shows that overall we see a reliable relationship between ERF and GMST trends in the CMIP5 ensemble, even if the ERF trend is negative.

We make the connection to multiple linear regression by writing each quantity as

$$x = \bar{x} + x'$$

where the overbar marks the ensemble average and the prime the across-ensemble variation. Linear expansion of equation (3) thus produces

$$\begin{aligned} \Delta \bar{T} + \Delta T' &= \frac{\Delta \bar{F}}{\bar{\alpha} + \bar{\kappa}} + \frac{1}{\bar{\alpha} + \bar{\kappa}} \Delta F' - \frac{\Delta \bar{F}}{(\bar{\alpha} + \bar{\kappa})^2} \alpha' \\ &\quad - \frac{\Delta \bar{F}}{(\bar{\alpha} + \bar{\kappa})^2} \kappa' + \varepsilon \end{aligned}$$

This equation holds for each start year separately and suggests the regression model

$$\Delta T'_j = \beta_0 + \beta_1 \Delta F'_j + \beta_2 \alpha'_j + \beta_3 \kappa'_j + \varepsilon_j, \quad j = 1, \dots, 75$$

We thus perform for each start year a multiple linear regression of $\Delta T'$ against $\Delta F'$, α' and κ' . The regression residual ε is interpreted as the contribution from internal variability. The complete regression-based prediction for GMST trend is obtained by adding the ensemble-mean trend to the regression for the across-ensemble variations:

$$\Delta \hat{T}_{\text{reg},j} = \Delta \bar{T} + \hat{\beta}_1 \Delta F'_j + \hat{\beta}_2 \alpha'_j + \hat{\beta}_3 \kappa'_j, \quad j = 1, \dots, 75 \quad (4)$$

where the caret marks the regression estimate. We note that for a model that has multiple realizations, the same $\Delta F'_j$, α'_j or κ'_j value is counted multiple times. The regression is performed separately for each period length over which trends are computed. We interpret the ensemble spread of the regression result $\Delta \hat{T}_{\text{reg},j}$, $j = 1, \dots, 75$, as the deterministic spread and the spread $\hat{\varepsilon}_j$, $j = 1, \dots, 75$, of the residuals as the quasi-random spread.

Deterministic versus quasi-random spread

For 15-year GMST trends, deterministic across-ensemble variations are smaller than internal variability, as shown by the comparison of the regression-based ensemble spread with the regression residuals (Fig. 2b and Fig. 2c, respectively). The regression result shows substantial time dependence in ensemble spread only for 15-year periods influenced by major volcanic eruptions, in particular the Mount Agung eruption in

1963 (Fig. 2b; the deterministic ensemble spread is particularly large in these periods; see Extended Data Fig. 4a). The distribution of residuals shows little time dependence, as evidenced by spread that is similar for all start years (Fig. 2c–f). The generally weak time dependence of the spread suggests that we can estimate the magnitudes of deterministic spread and internal variability from the marginal distributions obtained by time-averaging the distributions shown in Fig. 2b and Fig. 2c, respectively. The 5–95% range is $0.11^\circ \text{C per decade}$ for the regression result and $0.26^\circ \text{C per decade}$ for the residuals; internal variability thus dominates deterministic spread by a factor of 2.5. The dominance of internal variability in the ensemble spread of the 15-year GMST trends indicates that, viewed over the entire period 1900–2012, no systematic model error needs to be invoked when trying to explain differences between simulated and observed trends. In particular, the GMST spread due to feedback α is not systematically larger than the spread from either

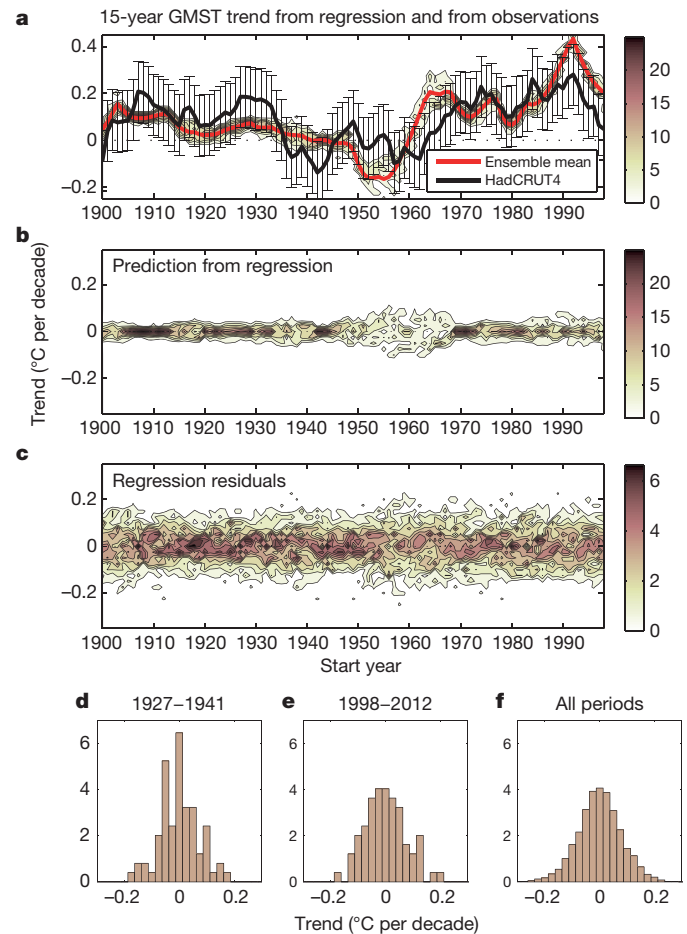


Figure 2 | Regression-based and observed 15-year GMST trends since 1900.

a, Joint relative frequency distribution of regression-based GMST trends (equation (4)) as a function of start year and trend size (in bins of $0.025^\circ \text{C per decade}$, as shown by colour scale), based on the reduced 75-member ensemble for which forcing information is available. The thick red line marks the ensemble average, the thick black line marks the observed trend and whiskers indicate the 5–95% confidence range derived from **f**. **b**, Joint relative frequency distribution of regression result (equation (4)) minus the ensemble-mean trend) as a function of start year and trend size (in bins of $0.025^\circ \text{C per decade}$). The P value of the regression has a median across start years of 0.075, based on the null hypothesis that all regression coefficients are zero. **c**, Joint relative frequency distribution of regression residual as a function of start year and trend size (in bins of $0.025^\circ \text{C per decade}$). **d**, Vertical cross-section of **c** for start year 1927. **e**, Vertical cross-section of **c** for start year 1998. **f**, Marginal distribution of regression residual as a function of trend size, obtained by time-averaging the joint distribution in **c**. All histograms are normalized such that their area integral is unity. In **a**–**c**, each vertical cross section is normalized, and the ordinate ranges are identical.

the ERF trend or ocean heat uptake efficiency, and is much smaller than the internal variability (Extended Data Fig. 4 and Fig. 2; see also ref. 12).

For any given start year, the residual spread is very similar to the full ensemble spread, implying that we can indeed use the ensemble spread as a measure of internal variability (compare Fig. 1b, c with Fig. 2d, e). Furthermore, identifying the ensemble spread of the regression residuals with internal variability allows us to characterize the component of observational uncertainty that arises from internal variability (Fig. 2a, f). This uncertainty does not concern the construction of the global average from individual station data (which has much smaller uncertainty⁵) but relates to the question of whether an observed trend is statistically significant (detectable) given serial correlation arising from internal variability¹⁸. Our model-based estimate of 0.26 °C per decade for the 5–95% confidence interval for observed 15-year GMST trends is slightly larger than the AR5 serial-correlation-based estimate for the uncertainty in the observed GMST trend over the hiatus period (0.2 °C per decade; ref. 4). We deem this an acceptable agreement given that the estimates were obtained through completely different approaches. We further note that the CMIP5 ensemble has been assessed to be generally consistent with observed historical decadal variability in GMST⁵, although on average it somewhat overestimates the global variability in the lower troposphere⁴⁵.

For most of the historical period, the entire ensemble of regression-based simulated 15-year GMST trends lies within the model-estimated 5–95% confidence interval of the observations (Fig. 2a). The regression-based simulated ensemble partly falls outside this interval during the cooling following the Mount Agung eruption and the subsequent warming recovery, as well as for start dates after 1990, which include the warming recovery following the Mount Pinatubo eruption and the surface warming hiatus (Fig. 2a). Because the phases of volcanically driven cooling and subsequent warming coincide with larger regression spread due to the ERF trend (Extended Data Fig. 4), we speculate that the implementation of volcanic forcing requires improvement in some climate models.

The ensemble spread of 62-year GMST trends is dominated by internal variability for start years early in the twentieth century, but for start years from 1910 onwards, the deterministic spread increases and dominates for start years 1920 and later (Fig. 3). The 5–95% range of the regression residuals is 0.059 °C per decade, compared with a deterministic range of 0.032 °C per decade for start year 1900 and 0.093 °C per decade for start year 1951. The 5–95% deterministic range for all 62-year trends is 0.081 °C per decade, which is larger by one-third than the 5–95% range from internal variability. Nevertheless, we see a substantial influence of internal variability even for GMST trends over 62 years.

When observational uncertainty is accounted for—again on the basis of the 5–95% confidence interval derived from quasi-random model spread—the ensemble-mean simulated 62-year GMST trend is consistent with the observed trend for all start years after around 1915; before that, the simulations tend to warm too little (Fig. 3a). After around 1945, the ensemble-mean simulated 62-year trend lies above the observed trend, although their difference is smaller than the range of internal variability. From around 1925 onward, both the largest and the smallest individual regression-based simulated trends lie outside the range defined by observations plus internal variability and are hence judged to be inconsistent with observations (Fig. 3a).

The cause of this inconsistency can be traced almost entirely to the contribution to the regression by the ERF trend (Fig. 3). By contrast, the magnitude of the contributions from α and κ is around 0.01 °C per decade or less for all start years (Fig. 3e, f). The deterministic ensemble spread in 62-year GMST trend is hence dominated by the spread in ERF throughout the twentieth century (Fig. 3).

Discussion

Viewed over the entire period since 1900, the differences between simulated and observed 15-year trends in GMST are dominated by internal

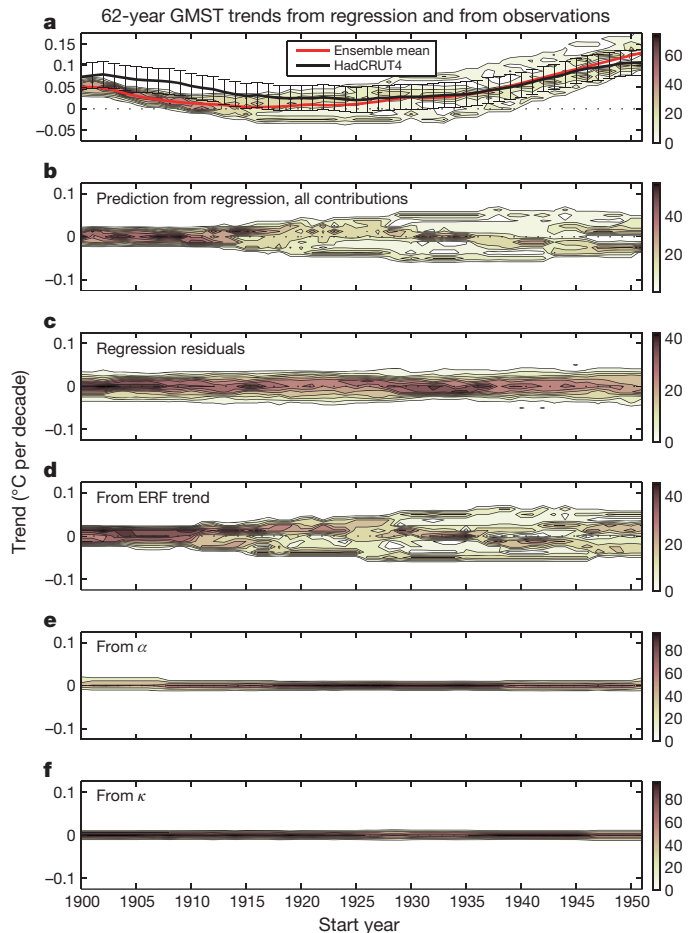


Figure 3 | Regression-based and observed 62-year GMST trends since 1900.

a, Joint relative frequency distribution of regression-based GMST trends (equation (4); shown by colour scale) as a function of start year and trend size, based on the reduced 75-member ensemble for which forcing information is available. The thick red line marks the ensemble average, the thick black line marks the observed trend and whiskers indicate the 5–95% confidence range derived from the marginal distribution of **c**. **b**, Joint relative frequency distribution of regression result (equation (4) minus the ensemble-mean trend) as a function of start year and trend size. All *P* values of the regression are below 0.001, based on the null hypothesis that all regression coefficients are zero. **c**, Joint relative frequency distribution of regression residual as a function of start year and trend size. **d**, Joint relative frequency distribution of regression contribution from trend in effective radiative forcing. **e**, Joint relative frequency distribution of regression contribution from climate feedback parameter α . **f**, Joint relative frequency distribution of regression contribution from ocean heat uptake efficiency κ . In all joint relative frequency distributions, GMST trend is collected in bins of 0.0125 °C per decade, and each vertical cross section is normalized such that its area integral is unity. All ordinate ranges are identical.

variability and hence arise largely by coincidence, with a minor contribution from volcanic forcing that is sometimes too strong in some models (Fig. 2). Furthermore, we confirm, and extend to all 15-year radiative forcing trends since 1900, the AR5 assessment for the hiatus period⁵ that the CMIP5 models show little systematic bias when compared with the AR5 best-estimate radiative forcing trend⁴⁶—despite the substantial scatter about the ensemble mean (Extended Data Fig. 2).

The generally dominant role of internal variability in shaping simulated 15-year GMST trends implies that internal variability also dominates the difference between simulations and observations during the hiatus period. This conclusion considerably sharpens the relative roles of internal variability, forcing error and response error, compared with the corresponding AR5 assessment⁵. Although there is no obvious contribution of forcing bias in the CMIP5 models (Extended Data Fig. 2),

the diagnosed radiative forcing is uncertain³⁵. Hence, our analysis cannot rule out a small contribution from a systematic forcing bias^{12,15,16,46–48} in the models. In particular, volcanic forcing is estimated to contribute to the difference between simulations and observations by up to 15% over 1998–2012¹², with large uncertainty in the magnitude. This is a contribution that our method cannot detect. Furthermore, the period 1998–2012 stands out as the only one during which the HadCRUT4 15-year GMST trend falls entirely outside the CMIP5 ensemble (if only narrowly), suggesting that the CMIP5 models could be missing a cooling contribution from the radiative forcing during the hiatus period^{12,15,16,46–48}, or that there has been an unusual enhancement of ocean heat uptake not simulated by any model¹⁹.

For 62-year GMST trends since 1900, the difference between simulations and observations is dominated by the spread in the radiative forcing trend in the models, with a smaller yet substantial influence of internal variability (Fig. 3). Our simple regression-based estimate of internal variability in 62-year GMST trends corresponds to a 17–83% range of ± 0.11 °C for the temperature change over six decades, which is in excellent agreement with the value of ± 0.10 °C that has been found for the period 1951–2010 using much more sophisticated formal methods of detection and attribution¹⁸.

There is scientific, political and public debate regarding the question of whether the GMST difference between simulations and observations during the hiatus period might be a sign of an equilibrium model response to a given radiative forcing that is systematically too strong, or, equivalently, of a simulated climate feedback α that is systematically too small (equation (2)). By contrast, we find no substantive physical or statistical connection between simulated climate feedback and simulated GMST trends over the hiatus or any other period, for either 15- or 62-year trends (Figs 2 and 3 and Extended Data Fig. 4). The role of simulated climate feedback in explaining the difference between simulations and observations is hence minor or even negligible. By implication, the comparison of simulated and observed GMST trends does not permit inference about which magnitude of simulated climate feedback—ranging from 0.6 to 1.8 W m⁻² °C⁻¹ in the CMIP5 ensemble—better fits the observations. Because observed GMST trends do not allow us to distinguish between simulated climate feedbacks that vary by a factor of three, the claim that climate models systematically overestimate the GMST response to radiative forcing from increasing greenhouse gas concentrations seems to be unfounded.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 August; accepted 26 November 2014.

- Knight, J. *et al.* Do global temperature trends over the last decade falsify climate predictions? *Bull. Am. Meteorol. Soc.* **90**, S22–S23 (2009).
- Wang, S. *et al.* Does the global warming pause in the last decade: 1999–2008? *Adv. Clim. Change Res.* **1**, 49–54 (2010).
- Liebmann, B., Dole, R. M., Jones, C., Blade, I. & Allured, D. Influence of choice of time period on global surface temperature trend estimates. *Bull. Am. Meteorol. Soc.* **91**, 1485–1491 (2010).
- Hartmann, D. L. *et al.* in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. *et al.*) 159–254 (Cambridge Univ. Press, 2013).
- Flato, G. *et al.* in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. *et al.*) 741–866 (Cambridge Univ. Press, 2013).
- Meehl, G. A., Arblaster, J. M., Fasullo, J. T., Hu, A. & Trenberth, K. E. Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nature Clim. Change* **1**, 360–364 (2011).
- Meehl, G. A. & Teng, H. Case studies for initialized decadal hindcasts and predictions for the Pacific region. *Geophys. Res. Lett.* **39**, L22705 (2012).
- Meehl, G. A., Hu, A., Arblaster, J. M., Fasullo, J. & Trenberth, K. E. Externally forced and internally generated decadal climate variability associated with the interdecadal Pacific oscillation. *J. Clim.* **26**, 7298–7310 (2013).
- Doblas-Reyes, F. J. *et al.* Initialized near-term regional climate change prediction. *Nature Commun.* **4**, 1715 (2013).
- Guemas, V., Doblas-Reyes, F. J., Andreu-Burillo, I. & Asif, M. Retrospective prediction of the global warming slowdown in the past decade. *Nature Clim. Change* **3**, 649–653 (2013).
- Fyfe, J. C., Gillett, N. P. & Zwiers, F. W. Overestimated global warming over the past 20 years. *Nature Clim. Change* **3**, 767–769 (2013).
- Santer, B. D. *et al.* Volcanic contribution to decadal changes in tropospheric temperature. *Nature Geosci.* **7**, 185–189 (2014).
- Risbey, J. S. *et al.* Well-estimated global surface warming in climate projections selected for ENSO phase. *Nature Clim. Change* **4**, 835–840 (2014).
- Huber, M. & Knutti, R. Natural variability, radiative forcing and climate response in the recent hiatus reconciled. *Nature Geosci.* **7**, 651–656 (2014).
- Solomon, S. *et al.* The persistently variable “background” stratospheric aerosol layer and global climate change. *Science* **333**, 866–870 (2011).
- Schmidt, G. A., Shindell, D. T. & Tsigaridis, K. Reconciling warming trends. *Nature Geosci.* **7**, 158–160 (2014).
- Stott, P., Good, P., Jones, G., Gillett, N. & Hawkins, E. The upper end of climate model temperature projections is inconsistent with past warming. *Environ. Res. Lett.* **8**, 014024 (2013).
- Bindoff, N. L. *et al.* in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. *et al.*) 867–952 (Cambridge Univ. Press, 2013).
- England, M. H. *et al.* Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nature Clim. Change* **4**, 222–227 (2014).
- Cohen, J. L., Furtado, J. C., Barlow, M., Alexeev, V. A. & Cherry, J. E. Asymmetric seasonal temperature trends. *Geophys. Res. Lett.* **39**, L04705 (2012).
- Kosaka, Y. & Xie, S.-P. Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature* **501**, 403–407 (2013).
- Lean, J. L. & Rind, D. H. How will Earth’s surface temperature change in future decades? *Geophys. Res. Lett.* **36**, L15708 (2009).
- Foster, G. & Rahmstorf, S. Global temperature evolution 1979–2010. *Environ. Res. Lett.* **6**, 044022 (2011).
- Kaufmann, R. K., Kauppi, H., Mann, M. L. & Stock, J. H. Reconciling anthropogenic climate change with observed temperature 1998–2008. *Proc. Natl Acad. Sci. USA* **108**, 11790–11793 (2011).
- Thompson, D. W. J., Wallace, J. M., Jones, P. D. & Kennedy, J. J. Identifying signatures of natural climate variability in time series of global-mean surface temperature: methodology and insights. *J. Clim.* **22**, 6120–6141 (2009).
- Santer, B. D. *et al.* Separating signal and noise in atmospheric temperature changes: the importance of timescale. *J. Geophys. Res.* **116**, D22105 (2011).
- Morice, C. P., Kennedy, J. J., Rayner, N. A. & Jones, P. D. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 data set. *J. Geophys. Res.* **117**, D08101 (2012).
- Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).
- Cowan, K. & Way, R. G. Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Q. J. R. Meteorol. Soc.* **140**, 1935–1944 (2014).
- Simmons, A. J., Willett, K. M., Jones, P. D., Thorne, P. W. & Dee, D. P. Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: inferences from reanalyses and monthly gridded observational data sets. *J. Geophys. Res.* **115**, D01110 (2010).
- Anderson, J. L. A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *J. Clim.* **9**, 1518–1530 (1996).
- Easterling, D. R. & Wehner, M. F. Is the climate warming or cooling? *Geophys. Res. Lett.* **36**, L08706 (2009).
- Stroeve, J. C. *et al.* Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. *Geophys. Res. Lett.* **39**, L16502 (2012).
- Notz, D., Haumann, F. A., Haak, H., Jungclauss, J. H. & Marotzke, J. Arctic sea-ice evolution as modeled by Max Planck Institute for Meteorology’s Earth system model. *J. Adv. Model. Earth Syst.* **5**, 173–194 (2013).
- Forster, P. M. *et al.* Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *J. Geophys. Res.* **118**, 1–12 (2013).
- Gregory, J. M. & Forster, P. M. Transient climate response estimated from radiative forcing and observed temperature change. *J. Geophys. Res.* **113**, D23105 (2008).
- Held, I. M. *et al.* Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *J. Clim.* **23**, 2418–2427 (2010).
- Andrews, T., Gregory, J. M., Webb, M. J. & Taylor, K. E. Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophys. Res. Lett.* **39**, L09712 (2012).
- Kuhlbrodt, T. & Gregory, J. M. Ocean heat uptake and its consequences for the magnitude of sea level rise and climate change. *Geophys. Res. Lett.* **39**, L18608 (2012).
- Vial, J., Dufresne, J.-L. & Bony, S. On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates. *Clim. Dyn.* **41**, 3339–3362 (2013).
- Gregory, J. M. *et al.* A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.* **31**, L03205 (2004).
- Block, K. & Mauritsen, T. Forcing and feedback in the MPI-ESM-LR coupled model under abruptly quadrupled CO₂. *J. Adv. Model. Earth Syst.* **5**, 676–691 (2013).
- Meraner, K., Mauritsen, T. & Voigt, A. Robust increase in equilibrium climate sensitivity under global warming. *Geophys. Res. Lett.* **40**, 5944–5948 (2013).
- Forster, P. *et al.* in *Climate Change 2007: The Physical Science Basis*. (eds Solomon, S. *et al.*) 129–234 (Cambridge Univ. Press, 2007).
- Santer, B. D. *et al.* Identifying human influences on atmospheric temperature. *Proc. Natl Acad. Sci. USA* **110**, 26–33 (2013).
- Myhre, G. *et al.* in *Climate Change 2013: The Physical Science Basis*. (eds Stocker, T. F. *et al.*) 659–740 (Cambridge Univ. Press, 2013).
- Fyfe, J. C., von Salzen, K., Cole, J. N. S., Gillett, N. P. & Vernier, J. P. Surface response to stratospheric aerosol changes in a coupled atmosphere-ocean model. *Geophys. Res. Lett.* **40**, 584–588 (2013).

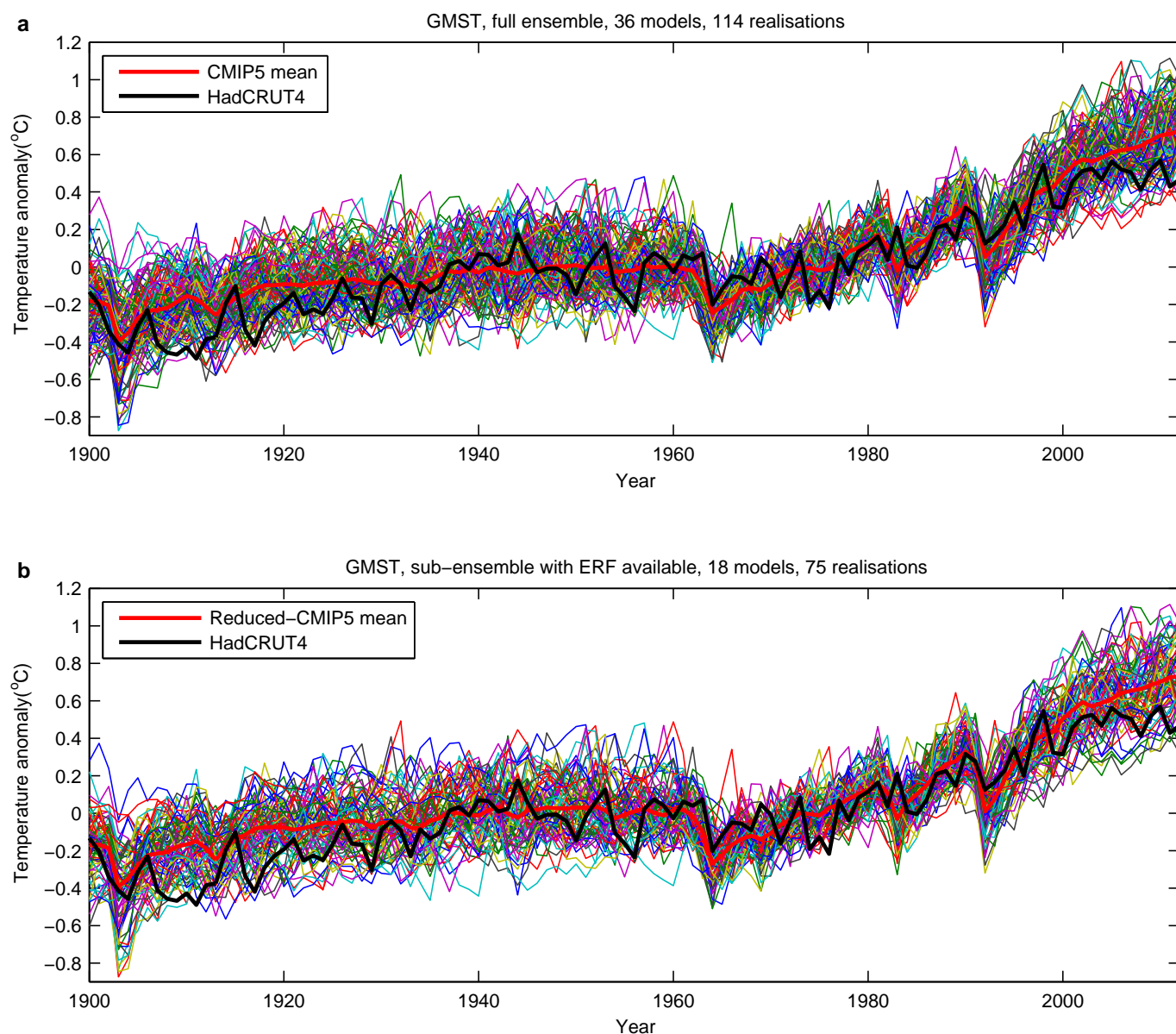
48. Ridley, D. A. *et al.* Total volcanic stratospheric aerosol optical depths and implications for global climate change. *Geophys. Res. Lett.* **41**, 7763–7769 (2014).

Acknowledgements We are indebted to J. Fyfe for making his CMIP5 GMST data set available to us, and to D. Notz, J. Risbey and B. Santer for comments on the manuscript. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups (names of models listed in Extended Data Table 1) for producing and making available their model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for

Earth System Science Portals. This work was supported by the Max Planck Society for the Advancement of Science (J.M.) and by a Royal Society Wolfson Merit Award and EPSRC grant EP/1014721/1 (P.M.F.).

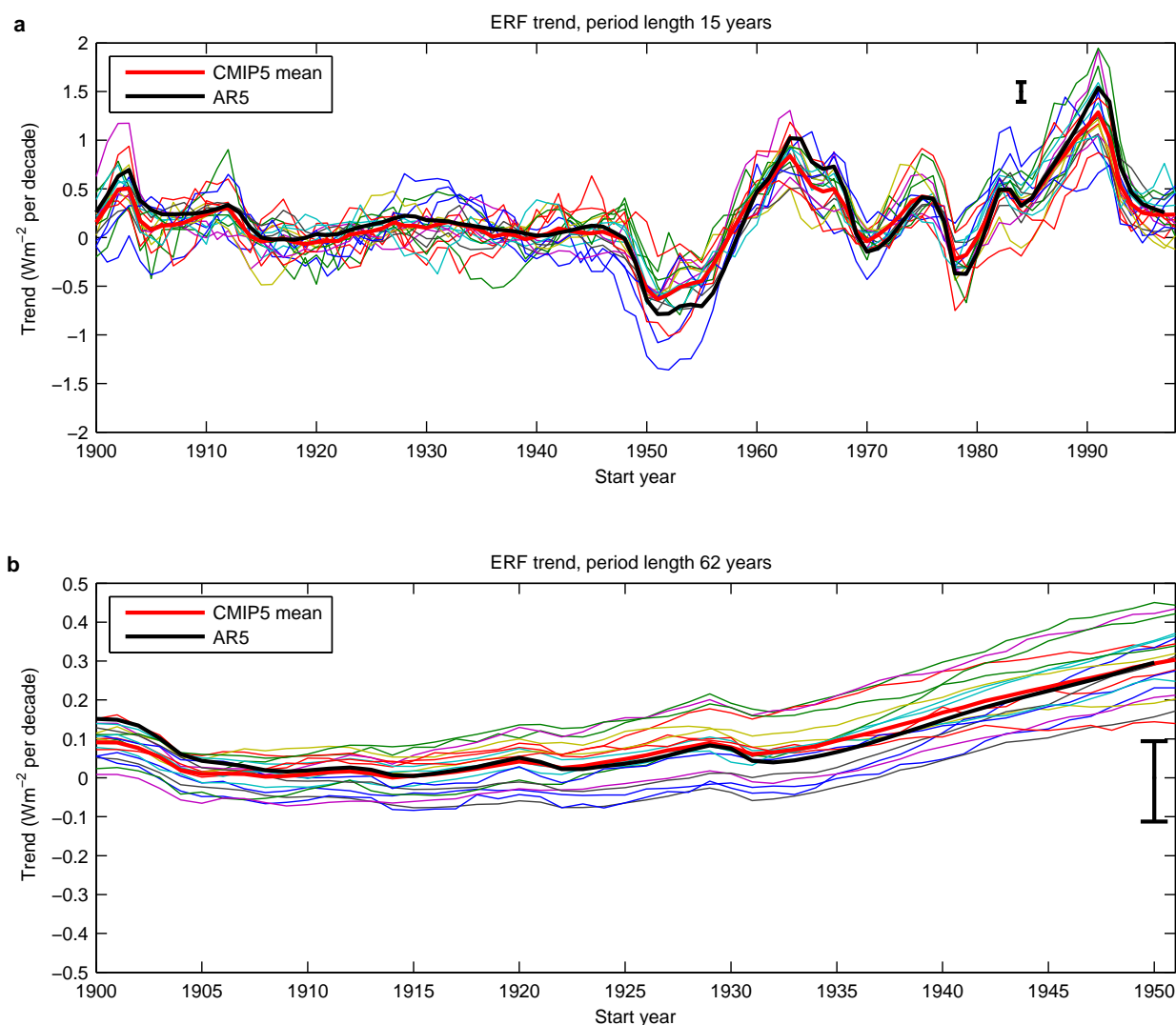
Author Contributions The authors jointly designed the study. J.M. analysed the data and wrote the manuscript. Both authors discussed the results and the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.M. (jochem.marotzke@mpimet.mpg.de).



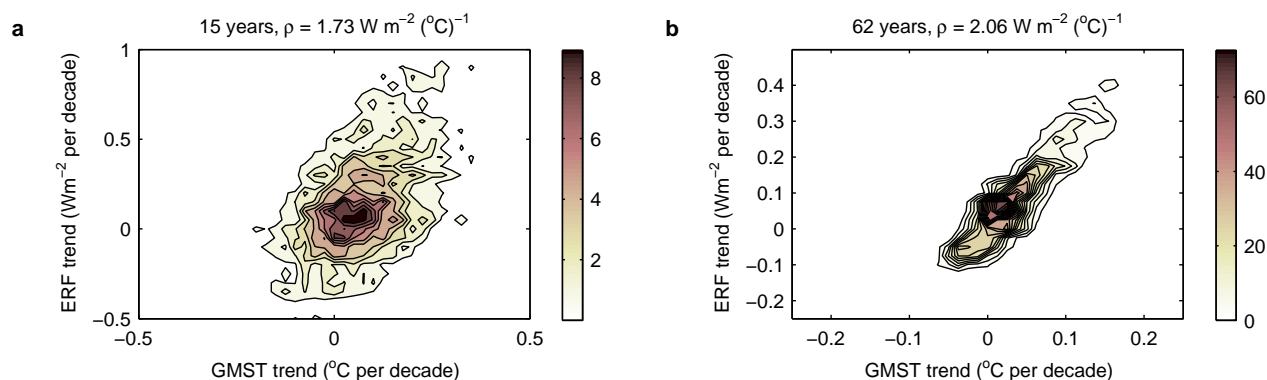
Extended Data Figure 1 | Observed and simulated time series of the anomalies in annually averaged GMST, from 1900 to 2012. All anomalies are differences from the 1961–1990 temporal mean of each individual time series. GMST is the globally averaged merged surface temperature (2 m height over land and surface temperature over the ocean). The figure shows single simulations for the CMIP5 models (thin lines), the multimodel ensemble mean

(thick red line) and the HadCRUT4²⁷ observations (thick black line). All model results have been subsampled using the HadCRUT4 observational data mask¹¹. **a**, 114 realizations from the CMIP5 archive, obtained with 36 different models. **b**, Subset of 75 realizations with the 18 different models for which information on ERF is available³⁵ (Extended Data Table 1). The two model ensembles are nearly indistinguishable.



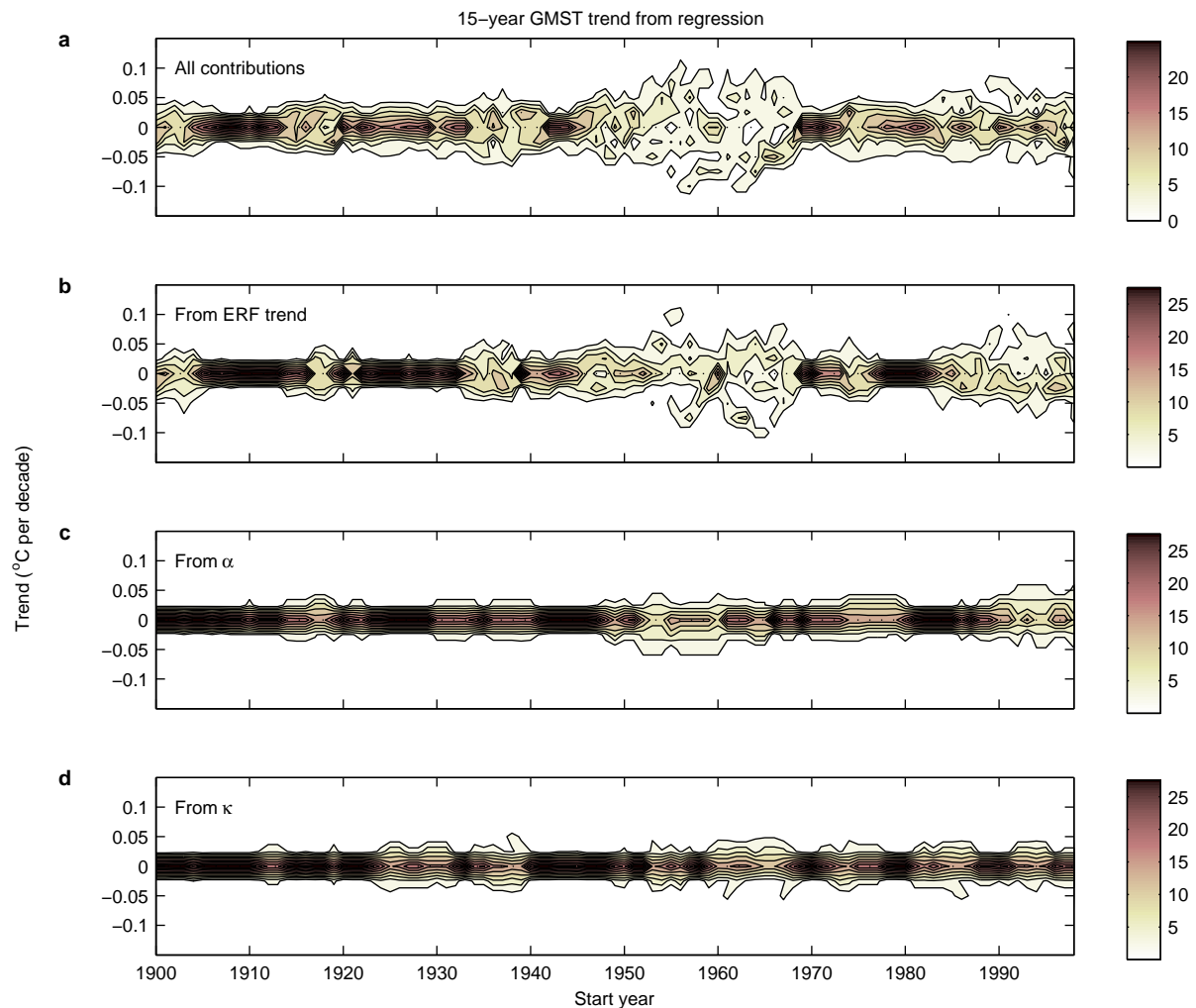
Extended Data Figure 2 | Time series of trends in ERF, as a function of start year. **a**, 15-year trends; **b**, 62-year trends. Thin coloured lines show individual models as diagnosed previously³⁵; if multiple realizations were available for a model, then the ensemble average of the individual diagnosed ERF time series for that model was given³⁵ and is shown here. The thick red line shows the ensemble average over all models. The thick black line shows the best estimate from AR5⁴⁶, including, for illustration, the 5–95% uncertainty range for the periods 1984–1998 (**a**) and 1951–2011 (**b**), taken from fig. 8.19 in ref. 46. These uncertainty ranges, both of which are around 0.2 W m^{-2} per decade, do not take into account observational biases such as those diagnosed in ref. 48. Despite the scatter of the CMIP5 ensemble trends, the ensemble mean is in good agreement with the AR5 best estimate for almost all start years. The AR5 best-estimate ERF sums time series of forcing across individual forcing terms. Individual time series of AR5 ERF were derived in different ways. Greenhouse gas concentrations (observed or inferred), stratospheric aerosol optical depth

and total solar irradiance were used to derive estimates of radiative forcing using simple formulae. Surface albedo forcing was derived from estimated anthropogenic vegetation trends. Ozone and aerosol forcings were derived from chemical transport model results with aspects of the forcing constrained by other modelling approaches or observations, or both. ERF sums rapid adjustments with traditional radiative forcings. Most time series in AR5 were based on traditional radiative forcings, and only CO_2 and aerosol forcings included an assessment of the rapid adjustment. In other cases ERF and radiative forcings were assumed to be the same. The AR5 ERF for the most recent 2000–2011 period included updated estimates of volcanic and solar forcing, taking into account the broader 2008–2009 solar minimum and post-2000 volcanic activity⁴⁶. These two cooling influences are not included in the CMIP5 ERF; it is hence surprising and unexplained why the CMIP5 ensemble-mean of 15-year ERF trends lies below the best-estimate AR5 ERF trend for the latest start years in **a**.



Extended Data Figure 3 | Joint relative frequency distribution as a function of GMST trend and ERF trend, for the reduced 75-member ensemble for which forcing information is available and all start years. a, 15-year trends; bin sizes are $0.025 \text{ °C per decade}$ and 0.05 W m^{-2} per decade for GMST and ERF trend, respectively. b, 62-year trends; bin sizes are 0.0125 °C per

decade and 0.025 W m^{-2} per decade for GMST and ERF trend, respectively. The 'climate resistance', ρ , is given by $\rho = \alpha + \kappa$ (refs 35–37). Each joint distribution is normalized such that its area integral is unity. Note the different axes, reflecting the much tighter correlation of the 62-year trends.



Extended Data Figure 4 | Regression-based 15-year GMST trends since 1900. **a**, Joint relative frequency distribution of regression result (equation (4) minus the ensemble-mean trend) as a function of start year and trend size. The P values of the regression have a median across start years of 0.075, based on the null hypothesis that all regression coefficients are zero. **b**, Joint relative frequency distribution of regression contribution from the trend in

ERF. **c**, Joint relative frequency distribution of regression contribution from the climate feedback parameter α . **d**, Joint relative frequency distribution of regression contribution from the ocean heat uptake efficiency κ . In all joint relative frequency distributions, GMST trend is collected in bins of 0.025°C per decade, and each vertical cross section is normalized such that its area integral is unity.

Extended Data Table 1 | CMIP5 models used in this study

Model name	Number of realisations	Forcing available?
ACCESS1-0	1	Y
ACCESS1-3	1	
bcc-csm1-1	3	Y
bcc-csm1-1-m	3	Y
BNU-ESM	1	
CanESM2	5	Y
CCSM4	6	Y
CESM1-BGC	1	
CESM1-CAM5	3	
CMCC-CM	1	
CMCC-CMS	1	
CNRM-CM5	10	Y
CSIRO-Mk3-6-0	10	Y
FIO-ESM	3	
GFDL-CM3	5	Y
GFDL-ESM2G	1	Y
GFDL-ESM2M	1	Y
GISS-E2-H	5	
GISS-E2-H-CC	1	
GISS-E2-R	6	Y
GISS-E2-R-CC	1	
HadCM3	10	
HadGEM2-AO	1	
HadGEM2-CC	1	
HadGEM2-ES	1	Y
IPSL-CM5A-LR	6	Y
IPSL-CM5A-MR	3	
IPSL-CM5B-LR	1	
MIROC5	5	Y
MIROC-ESM	3	Y
MIROC-ESM-	1	
MPI-ESM-LR	3	Y
MPI-ESM-MR	3	
MRI-CGCM3	3	Y
NorESM1-M	3	

The originating institutions and publications documenting the models are listed comprehensively in table 9.A1 of ref. 5.

An *Arabidopsis* gene regulatory network for secondary cell wall synthesis

M. Taylor–Teeple^{1,2*}, L. Lin^{3†*}, M. de Lucas^{1,2*}, G. Turco^{1,2}, T. W. Toal^{1,2}, A. Gaudinier^{1,2}, N. F. Young³, G. M. Trabucco³, M. T. Veling³, R. Lamothe³, P. P. Handakumbura³, G. Xiong⁴, C. Wang¹, J. Corwin⁵, A. Tsoukalas^{2,6}, L. Zhang⁷, D. Ware^{7,8}, M. Pauly⁴, D. J. Kliebenstein⁵, K. Dehesh¹, I. Tagkopoulos^{2,6}, G. Breton^{9†}, J. L. Pruneda–Paz⁹, S. E. Ahnert¹⁰, S. A. Kay^{9†}, S. P. Hazen³ & S. M. Brady^{1,2}

The plant cell wall is an important factor for determining cell shape, function and response to the environment. Secondary cell walls, such as those found in xylem, are composed of cellulose, hemicelluloses and lignin and account for the bulk of plant biomass. The coordination between transcriptional regulation of synthesis for each polymer is complex and vital to cell function. A regulatory hierarchy of developmental switches has been proposed, although the full complement of regulators remains unknown. Here we present a protein–DNA network between *Arabidopsis thaliana* transcription factors and secondary cell wall metabolic genes with gene expression regulated by a series of feed–forward loops. This model allowed us to develop and validate new hypotheses about secondary wall gene regulation under abiotic stress. Distinct stresses are able to perturb targeted genes to potentially promote functional adaptation. These interactions will serve as a foundation for understanding the regulation of a complex, integral plant component.

Plant cell shape and function are in large part determined by the cell wall. Almost all cells have a primary wall surrounding the plasma membrane. Specialized cell types differentiate by depositing a secondary cell wall upon cessation of cell elongation. In addition to providing mechanical support for water transport and a barrier against invading pathogens, the polymers contained within the wall are an important renewable resource for humans as dietary fibre, as raw material for paper and pulp manufacturing, and as a potential feedstock for biofuel production. Secondary cell walls account for the bulk of renewable plant biomass available globally.

The secondary cell wall consists of three types of polymer—cellulose, hemicelluloses and lignin—and is found in xylem, fibres and anther cells. Cellulose microfibrils form a main load-bearing network. Hemicelluloses include xylans, glucans, and mannans. Lignin is a complex phenylpropanoid polymer that imparts ‘waterproofing’ capacity as well as mechanical strength, rigidity and environmental protection. Despite the importance of the plant secondary cell wall, our knowledge of the precise regulatory mechanisms that give rise to these metabolites is limited. The expression of cell wall-associated genes is tightly spatiotemporally co-regulated^{1,2}. However, the pervasive functional redundancy within transcription factor families, the combinatorial complexity of regulation, and activity in a small number of cell types render functional characterization from single gene experiments difficult. A model of master regulators has been proposed with NAC domain and homeobox HD–ZIP Class III (HD–ZIPIII) transcription factors initiating cell specification and secondary cell wall synthesis in *Arabidopsis thaliana*. In this model, VASCULAR-RELATED NAC DOMAIN6 (VND6) and VND7 are sufficient but not necessary to regulate xylem vessel formation; additionally, the HD–ZIPIII transcription factor PHABULOSA

(PHB) also regulates vessel formation, and acts in a highly redundant manner with four other HD–ZIPIII factors³. In anthers, two NAC domain transcription factors, NAC SECONDARY WALL THICKENING1 (NST1) and NST2, are sufficient to drive the secondary cell wall biosynthetic program, but act redundantly⁴. Thus, regulation of this process is highly redundant and combinatorial. However, no comprehensive map of interactions has been developed at cell-type-resolution over time, nor have upstream regulators been identified. We therefore chose to pursue a network-based approach to comprehensively characterize the transcriptional regulation of secondary cell wall biosynthesis.

Mapping the secondary cell wall synthesis regulatory network

To systematically map this regulatory network at cell-type-resolution, we used a combination of high-spatial-resolution gene expression data⁵ and the literature^{1,6} to identify fifty genes implicated in xylem cell specification. These included transcription factors and enzymes involved in cellulose, hemicellulose and lignin biosynthesis that are expressed in root xylem cells (Supplementary Table 1; Methods). Selection of both developmental regulators and downstream functional genes allowed us to interrogate upstream regulatory events that determine xylem specification and differentiation associated with secondary cell wall synthesis. Promoter sequences were screened using an enhanced yeast one hybrid (Y1H) assay against 467 (89%) of root-xylem-expressed transcription factors⁷. Protein interactions were identified for 45 of the promoters (Supplementary Table 2). The final network comprises 242 genes and 617 protein–DNA interactions (Fig. 1a; http://gturco.github.io/trenzalore/stress_network). Thirteen of the transcription factors have been previously identified as having a role in xylem development or

¹Department of Plant Biology, University of California Davis, One Shields Avenue, Davis, California 95616, USA. ²Genome Center, University of California Davis, One Shields Avenue, Davis, California 95616, USA. ³Biology Department, University of Massachusetts, Amherst, Massachusetts 01003, USA. ⁴Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, California 94720, USA. ⁵Department of Plant Sciences, University of California Davis, One Shields Avenue, Davis, California 95616, USA. ⁶Department of Computer Science, University of California Davis, One Shields Avenue, Davis, California 95616, USA. ⁷Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA. ⁸US Department of Agriculture, Agricultural Research Service, Ithaca, New York 14853, USA. ⁹Section of Cell and Developmental Biology, Division of Biological Sciences, University of California San Diego, La Jolla, California 92093, USA. ¹⁰Theory of Condensed Matter Group, Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, UK. †Present addresses: Division of Plant Sciences, University of Missouri, Columbia, Missouri 65211, USA (L.L.); Department of Integrative Biology and Pharmacology, University of Texas Health Science Center at Houston, Houston, Texas 77030, USA (G.B.); Molecular and Computational Biology Section, University of Southern California, Los Angeles, California 90089, USA (S.A.K.).

*These authors contributed equally to this work.

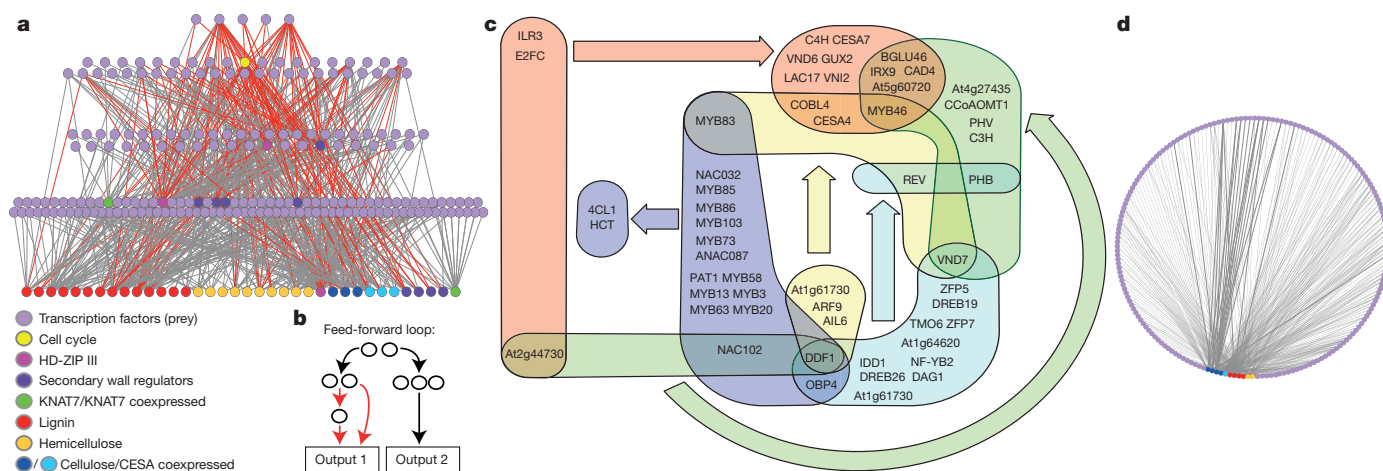


Figure 1 | Regulators of xylem development and secondary cell wall biosynthesis. **a**, Gene regulatory network for secondary cell wall biosynthesis in *Arabidopsis* root xylem. Nodes, transcription factors or promoters; edges,

secondary cell wall biosynthesis. Six of the transcription factors were previously shown to bind to these promoters and a further nine of the protein–DNA interactions were implied in gene expression studies, that is, without demonstrating direct binding^{6,8–11}. These interactions represent independent validation of our approach (Supplementary Table 2; Extended Data Fig. 1). All together, the network contains 601 novel interactions, although false negatives and false positives are a component of all network approaches¹².

Our Y1H approach revealed a highly interconnected regulatory network. On average, each cell wall gene promoter was bound by 5 transcription factors from 35 protein families with over-representation of AP2-EREBP, bHLH, C2H2, C2C2-GATA and GRAS gene families (Supplementary Table 3). Our network now adds an additional layer of gene regulation with novel factors upstream of *VND6* and *VND7* and supports feed-forward loops^{9,11,13} as an overarching theme for regulation of this developmental process with a total of 96 such loops (Fig. 1a, b).

To organize the network, we employed a power graph compression approach to condense the network into overlapping node sets with similar connectivity. Protein–DNA interactions (edges) between proteins and promoters (nodes) in the original network were replaced by ‘power edges’ between overlapping ‘power nodes’¹⁴. A power edge exists between suites of transcription factors that bind to the same set of promoters. Using this approach, 24 power edges were observed (Supplementary Table 4; Fig. 1c). Some sets could be distinguished on the basis of target gene function. For instance, one power edge connects 16 transcription factors with promoters of two lignin genes, *4CL1* and *HCT*, while another power edge connects three transcription factors with genes related to cellulose and hemicellulose biosynthesis such as *CESA4*, *CESA7*, *IRX9*, *COBL4* and *GUX2*.

Testing interactions predicted by the network

Using our network, we hypothesized that E2Fc is a key upstream regulator of *VND6*, *VND7* and secondary cell wall biosynthesis genes. This hypothesis is based on our finding that E2Fc bound to 23 promoters including those of *VND6*, *VND7* and *MYB46*, and cellulose-, hemicellulose- and lignin-associated genes (Fig. 2a). *VND7* and *MYB46* are also known to bind to the promoters of many of these genes as well^{9,13,15}, creating a suite of feed-forward loops. E2Fc is a known negative regulator of endoreduplication^{16,17}. Before terminally differentiating, xylem cells elongate and likely undergo endoreduplication before secondary cell wall deposition. E2Fc can act as a transcriptional repressor^{16–18} as well as a transcriptional activator^{19–22} and here we report both. E2Fc activated *VND7* expression in a dose-dependent manner (Fig. 2b and Extended Data Fig. 2a, b) in transient assays, but not in the presence of RETINOBLASTOMA-RELATED (RBR) protein, as is typical of E2F

protein–DNA interactions. Edges in feed-forward loops are red. **b**, A sample feed-forward loop in red. **c**, ‘Power edges’ between node sets. **d**, The secondary wall network from sub-fragments of cell wall promoters.

transcription factors (Extended Data Fig. 2c). In an *E2Fc*-overexpressor line with the amino terminus deleted to overcome post-translational degradation^{16,17}, regulation of *VND7* expression by extremely high or

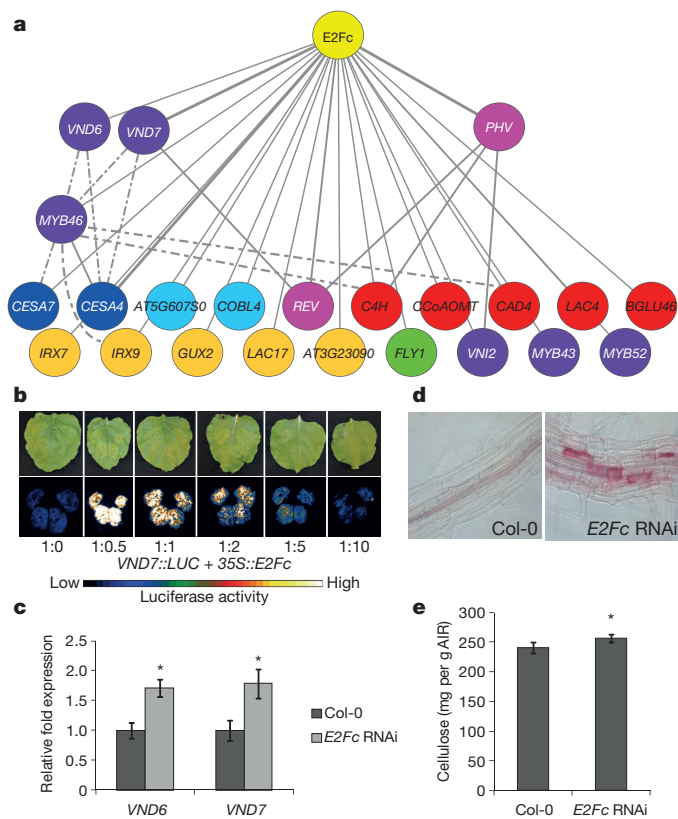


Figure 2 | E2Fc represses secondary cell wall gene biosynthesis. **a**, E2Fc-DNA interactions. Solid edges, Y1H; dashed edges, literature. **b**, Bright field (top) and dark-field (bottom) of representative leaves ($n = 20$) expressing *VND7::LUC* or together with *35S::E2Fc* in 1:0.1, 1:1, 1:2, 1:5, and 1:10 ratios, respectively. **c**, *VND6* and *VND7* expression relative to *UBC10* control in an *E2Fc* RNA interference (RNAi) line relative to wild type. $n = 2$ biological replicates with 3 technical replicates. **d**, **e**, Phloroglucinol staining of lignin ($n = 6$ per genotype (Col-0 and *E2Fc* RNAi, respectively), representative images shown) (**d**) and crystalline cellulose in wild-type and *E2Fc*-knockdown roots ($n = 3$ pooled samples for each genotype, each pooled sample has approximately 1,000 individuals) (**e**). AIR, alcohol-insoluble residues. For all panels, $*P < 0.05$ from Student's *t*-test and data are means \pm s.d.

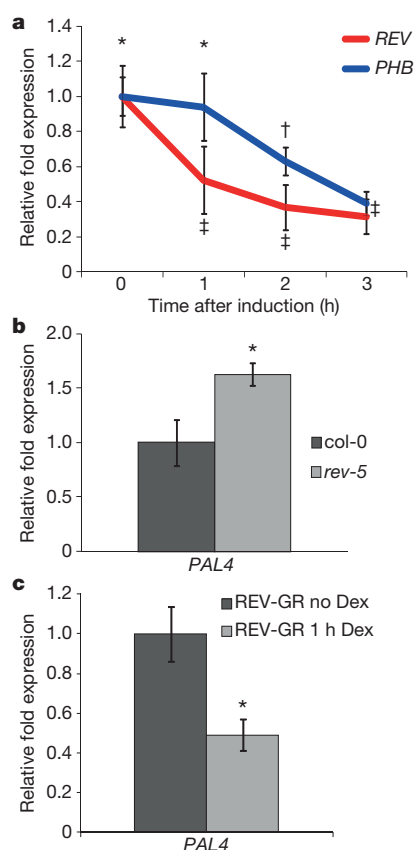


Figure 3 | Tissue-specific VND7 regulation and VND7 targets. **a**, *REV* and *PHB* expression relative to β -tubulin control following dexamethasone treatment of 35S::VND7:VP16:GR relative to untreated. $n = 3$ biological replicates; * significantly different from †, ‡ significantly different from ‡, and ‡ significantly different from ‡, $P < 0.01$. **b**, *PAL4* expression relative to *AT5G15710* control in *rev-5* relative to wild-type. **c**, *PAL4* expression relative to *UBC21* control following one hour dexamethasone (Dex) treatment of 35S::REV:GR relative to untreated. * $P < 0.05$ for panels **b** and **c**, $n = 2$ biological replicates with 3 technical replicates. All panels show data as means \pm s.d., with P calculated from Student's t -test.

low *E2Fc* levels resulted in *VND7* repression, whereas moderate *E2Fc* levels resulted in *VND7* activation (Extended Data Fig. 2b). This dynamic regulation was also observed in an *E2Fc*-knockdown line²³, where transcript abundance of *VND6* and *VND7* was significantly increased (Fig. 2c). Based on our results, we propose that *E2Fc* acts in a complex, concentration-dependent manner to regulate gene expression either as an activator or a repressor. Coincident with the repression observed in *E2Fc*-knockdown lines, ectopic patches of lignin were observed near the root–shoot junction using phloroglucinol staining (Fig. 2d). A significant increase in crystalline cellulose in the knockdown line was observed using an Updegraff assay (Fig. 2e).

All five HD-ZIPIII transcription factors, including REVOLUTA (*REV*), PHB, and PHAVOLUTA are jointly necessary for xylem cell specification and secondary wall synthesis³. We found that *VND7* bound *REV* and *PHB* promoters in yeast. *VND7* has been shown to act as a transcriptional activator⁹ or as a repressor when complexed with VNI2²⁴. With a dexamethasone-inducible version of *VND7*²⁵, transcript levels of *REV* and *PHB* were significantly decreased by 2.5-fold following induction (Fig. 3a). The *REV* transcription factor bound to the promoter of the lignin biosynthesis gene *PHENYLALANINE AMMONIA LYASE4* (*PAL4*). In a *rev-5* loss-of-function mutant, *PAL4* significantly increased in transcript abundance (Fig. 3b) and transient induction of *REV* by a glucocorticoid receptor fusion²⁶ resulted in a decrease of *PAL4* expression (Fig. 3c). Taken together, these data suggest that *E2Fc* can activate *VND7* expression in a dose-dependent manner, while *VND7*, possibly

in concert with VNI2, can repress *REV* expression, and *REV* can repress expression of *PAL4*. This series of interactions predicted by the network model and tested by perturbation analyses ensures that activation of *VND7* and coordination of lignin biosynthesis is tightly regulated.

We next sought to identify all transcription factors that potentially regulate secondary cell wall biosynthesis genes, not just in root xylem cells but also in above-ground cell types including xylary fibres, inter-fascicular fibres and anthers. Many of the biosynthetic genes downstream of the key NAC domain transcription factors act in both the root and the shoot⁹. To expand the network, we used Y1H to screen multiple smaller promoter fragments of a subset of promoters included in the root xylem network, including genes associated with cellulose, hemicellulose and lignin biosynthesis against a library of 1,664 full-length *Arabidopsis* transcription factors (Supplementary Tables 5, 6). We observed a total of 413 interactions that included proteins from 36 of the 75 protein families tested (Supplementary Table 7; Fig. 1d; http://gturco.github.io/trenzalore/secondary_cell_wall). We found an over-representation of AP2-EREBP, bZip, ZF-HD, MYB and GeBP families (Supplementary Table 8). Each promoter interacted with an average of 38 different proteins, generating even more possibilities for combinatorial, redundant or condition-specific gene regulation. Like the root xylem network, previously reported protein–DNA interactions were observed in this screen including MYB46 and MYB83 binding the promoters of *CESA* genes (Supplementary Table 7)^{8,27}. Since most of these interactions were novel, a subset was additionally validated. Transient expression of AIL1, MYB83, MYB54, NAC92, NST2 and SND1 caused a significant increase in *CESA4::LUC* activity in tobacco (measured by luciferase activity), indicating binding and activation of the *CESA4* promoter (Fig. 4a). We further tested three regions of the *CESA4* promoter with two NAC family proteins, SND1 and NST2 (Fig. 4b, c), using an *in vitro* electrophoretic mobility shift assay (EMSA). Extracts of *Escherichia coli* expressing either glutathione-S-transferase-conjugated NST2 (*GST::NST2*) or *GST::SND1* in the presence of a *CESA4-2pr* promoter probe produced DNA species with retarded mobility (Fig. 4b, c). We also observed binding of the *CESA7*, *CESA8* and *KOR* promoter fragments with the NST2 protein and *CESA8* with the SND1 protein (Extended Data Fig. 3). These interactions between NST2 and *CESA4*, *CESA8*, and *KOR* promoters were further confirmed *in planta* by chromatin immunoprecipitation (ChIP). An antibody to green fluorescent protein (GFP) was used to immunoprecipitate NST2 protein from extracts of 35S::NST2::GFP plants. The complex was significantly enriched for fragments from the *CESA4*, *CESA8* and *KOR* promoters (Fig. 4d). The tracheary element-regulating cis-element (TERE, CTTNAAAGCNA) is a direct target of VND6^{28,29}. A perfect TERE is present in the *CESA4* promoter (CTTGAAAGCTA) and TERE-like sequences are present in *CESA8* (CTTCAATGTTA) and *KOR* (CTTGAAAATGA). Taken together, these data clearly demonstrate that the expression of *CESA4* and other secondary cell wall genes is mediated by the direct binding of the NAC-domain binding transcription factors NST2 and SND1 to the target gene promoters via the TERE.

Abiotic stress can co-opt the xylem regulatory network

Having generated a gene regulatory network supported by *in vivo* and *in vitro* approaches, we sought to test if the model could allow us to predict responses under abiotic stress perturbation. Co-opting a developmental regulatory network is likely a key mechanism to facilitate adaptation in response to stress. Thus, we hypothesized that stress responses are likely integrated into the gene regulatory network that determines xylem cell specification and differentiation and that we can predict the exact genes that these stresses manipulate within our network.

We first identified genes within the network whose expression was altered specifically in the root vasculature in response to salt, sulphur, iron and pH stress^{30,31} and nitrogen influx³². Genes within the root xylem secondary cell wall network were significantly differentially regulated in response to sulphur stress, salt stress and iron deprivation (Supplementary Table 9). Substantial overlap was observed between iron

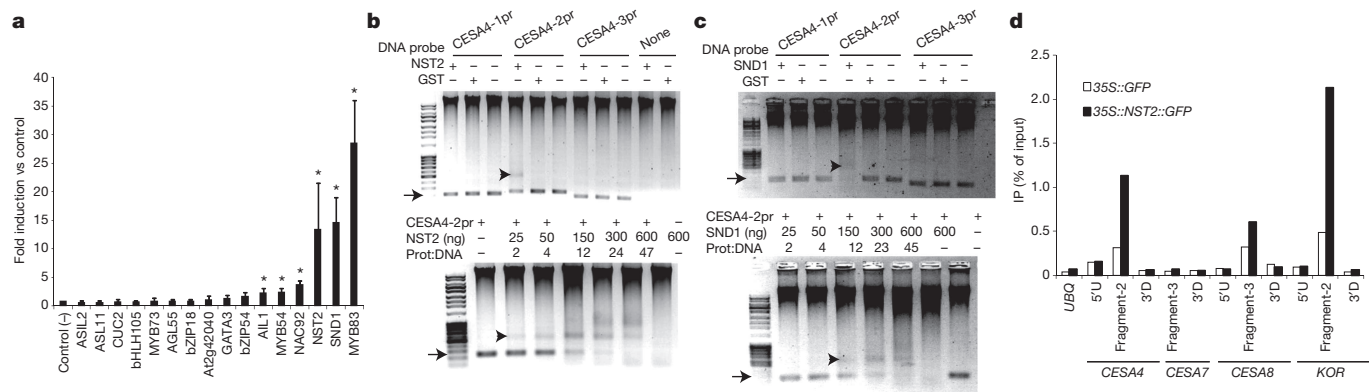


Figure 4 | Multiple transcription factors bind the *CESA4* promoter.

a, Activation of *CESA4::LUC* by transcription factors in tobacco ($n = 5$). * $P < 0.05$ based on Student's t -test. Data are means \pm s.d. **b**, **c**, EMSA with

deprivation and salt stress gene responses and was further characterized (Fig. 5a). We filtered the xylem network to include only genes differentially expressed in salt or iron, creating stress-specific sub-networks (Extended Data Fig. 4). Previously, we determined that key developmental transcription factors have significantly more upstream regulators compared to other genes³³. In response to iron deprivation, *REV* has the most upstream regulators, while in response to salt stress, *VND7* and *MYB46* have the most upstream regulators.

On the basis of these data from the iron-deprivation sub-network, we hypothesized that *REV* plays a key role in regulating secondary cell wall development in response to iron deprivation. To additionally determine directionality and sign (activation or repression) in the network, we constructed a network of 16 key nodes using the consensus network from four unsupervised and one supervised network inference method. *REV* was also predicted to be an important regulator of lignin biosynthesis gene expression in response to iron deprivation using these methods (Extended Data Fig. 5). First, to test the model-generated prediction that lignin biosynthetic gene expression is altered in response to iron deprivation, we measured phenylpropanoid-related gene expression. An increase in *4CL1*, *PAL4* and *HCT* gene expression was observed (Fig. 5b). Additionally, iron deprivation stress altered the timing and spatial distribution of the *4CL1* transcript (Fig. 5c). These expression changes are accompanied by an increase in fuchsin staining, indicative of increased phenylpropanoid deposition (Extended Data Fig. 6b). Expression in a *rev-5* loss-of-function mutant in iron-deficient conditions revealed a *REV*- and stress-dependent influence on *CCoAOMT1*, *PAL4* and *HCT* expression (Fig. 5d), thus validating our model predictions.

In the high-salinity sub-network *VND7* and *MYB46* contain the most upstream regulators (Extended Data Fig. 4). *VND7* and *MYB46* expression is greatly increased in roots in response to salt stress, but lignin biosynthetic gene expression is unaltered (Fig. 5e; Extended Data Fig. 6a). In corroboration with this hypothesis, the network model constructed using the described *in silico* methods also predicts *VND7* and *MYB46* as main regulators in response to salt stress but not iron deprivation (Extended Data Fig. 7), and indeed this was observed with an expansion of the domain of *VND7* expression after salt treatment but not iron deprivation (Fig. 5e, f; Extended Data Fig. 6c). In conjunction with this ectopic increase, we observed an additional strand of metaxylem in roots exposed to high salinity (Fig. 5g).

Discussion

Owing to functional redundancy among regulators of secondary cell wall biosynthesis, transcription factors have largely eluded identification by loss-of-function genetic screens. Our network approach has identified hundreds of novel regulators and provided considerable insight into the developmental regulation of xylem cell differentiation. The

NST2 (**b**) and *SND1* (**c**), with promoters. Arrowheads indicate protein–DNA complexes, arrows indicate free probe. **d**, ChIP of *NST2*–GFP with *CESA4*, *CESA7*, *CESA8* and *KOR* promoters. 3' D, 3' downstream; 5' U, 5' upstream.

network, which includes a cell cycle regulator, is comprised of many feed-forward loops that are likely to ensure robust regulation of this process (Fig. 5h). Accordingly, we revealed that perturbation at distinct

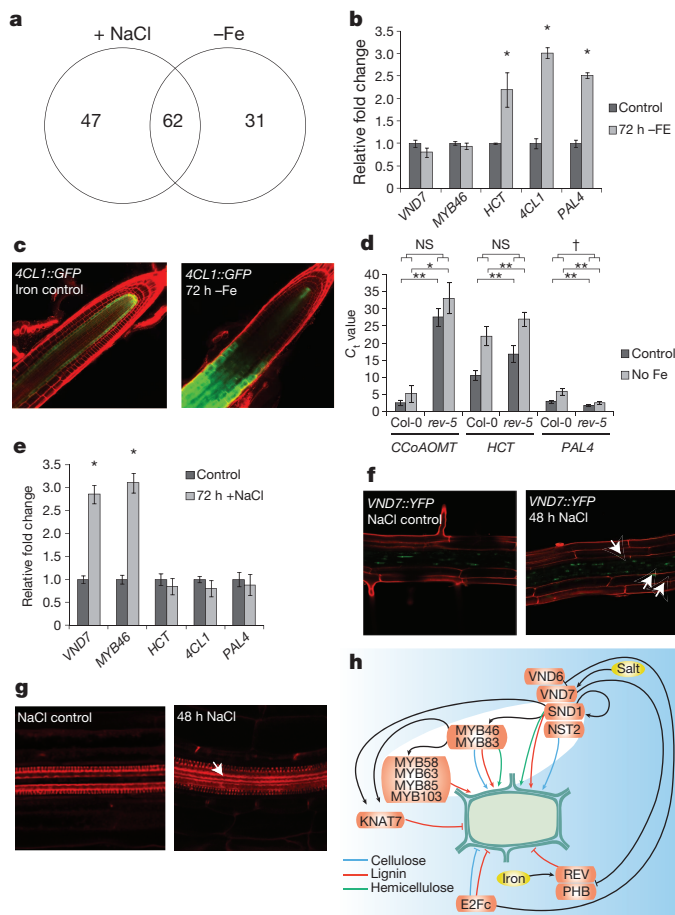


Figure 5 | The xylem-specific gene regulatory network is responsive to high salinity and iron deprivation.

a, Network genes responsive to high salinity and/or iron deprivation. **b**, *VND7*, *HCT*, *4CL1*, *PAL4* expression after iron deprivation. **c**, *4CL1::GFP* expression after iron deprivation (representative images shown, $n = 4$ per line). **d**, Lignin gene expression after iron deprivation in *rev-5*. * $P \leq 0.01$; ** $P \leq 0.001$; † $P \leq 0.0001$; P values from ANOVA. **e**, *VND7*, *HCT*, *4CL1*, *PAL4* expression after NaCl. **b**, **d**, **e**, Expression relative to *UBC10* and *PP2AA3* controls. $n = 2$ biological replicates with 3 technical replicates. **b**, **e**, * $P \leq 0.01$ based on Student's t -test and data are means \pm s.d. **f**, **g**, Representative images of *VND7::YFP* ($n = 5$) (**f**) and fuchsin-staining ($n = 5$) (**g**) after NaCl. Arrows, non-stele cells (**f**) and extra metaxylem strand (**g**). **h**, Proposed regulation of secondary wall biosynthesis.

nodes changes the network subtly, including phenylpropanoid biosynthesis in response to iron deprivation and ectopic xylem cell differentiation in response to salt stress (Fig. 5h). We anticipate that these findings will be instrumental in biotechnology and in our understanding of cell fate acquisition.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 November 2013; accepted 20 November 2014.

Published online 24 December 2014.

- Brown, D. M., Zeef, L. A. H., Ellis, J., Goodacre, R. & Turner, S. R. Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell* **17**, 2281–2295 (2005).
- Persson, S., Wei, H., Milne, J., Page, G. P. & Somerville, C. R. Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc. Natl Acad. Sci. USA* **102**, 8633–8638 (2005).
- Carlsbecker, A. *et al.* Cell signalling by microRNA165/6 directs gene dose-dependent root cell fate. *Nature* **465**, 316–321 (2010).
- Mitsuda, N., Seki, M., Shinozaki, K. & Ohme-Takagi, M. The NAC transcription factors NST1 and NST2 of *Arabidopsis* regulate secondary wall thickenings and are required for anther dehiscence. *Plant Cell* **17**, 2993–3006 (2005).
- Brady, S. M. *et al.* A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* **318**, 801–806 (2007).
- Zhong, R., Lee, C., Zhou, J., McCarthy, R. L. & Ye, Z.-H. A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell* **20**, 2763–2782 (2008).
- Gaudinier, A. *et al.* Enhanced Y1H assays for *Arabidopsis*. *Nature Methods* **8**, 1053–1055 (2011).
- Kim, W.-C., Ko, J.-H. & Han, K.-H. Identification of a *cis*-acting regulatory motif recognized by MYB46, a master transcriptional regulator of secondary wall biosynthesis. *Plant Mol. Biol.* **78**, 489–501 (2012).
- Yamaguchi, M. *et al.* VASCULAR-RELATED NAC-DOMAIN 7 directly regulates the expression of a broad range of genes for xylem vessel formation. *Plant J.* **66**, 579–590 (2011).
- Zhou, J., Lee, C., Zhong, R. & Ye, Z. H. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *Plant Cell* **21**, 248–266 (2009).
- Hussey, S. G., Mizrahi, E., Creux, N. M. & Myburg, A. A. Navigating the transcriptional roadmap regulating plant secondary cell wall deposition. *Front. Plant Sci.* **4**, 325 (2013).
- Walhout, A. J. M. What does biologically meaningful mean? A perspective on gene regulatory network validation. *Genome Biol.* **12**, 109 (2011).
- Kim, W.-C., Kim, J.-Y., Ko, J.-H., Kang, H. & Han, K.-H. Identification of direct targets of transcription factor MYB46 provides insights into the transcriptional regulation of secondary wall biosynthesis. *Plant Mol. Biol.* **85**, 589–599 (2014).
- Ahnert, S. E. Power graph compression reveals dominant relationships in genetic transcription networks. *Mol. Biosyst.* **9**, 2681–2685 (2013).
- Kim, W.-C. *et al.* MYB46 directly regulates the gene expression of secondary wall-associated cellulose synthases in *Arabidopsis*. *Plant J.* **73**, 26–36 (2013).
- del Pozo, J. C., Diaz-Trivino, S., Cisneros, N. & Gutierrez, C. The balance between cell division and endoreplication depends on E2FC-DPB, transcription factors regulated by the ubiquitin-SCF^{SKP2A} pathway in *Arabidopsis*. *Plant Cell* **18**, 2224–2235 (2006).
- del Pozo, J. C., Boniotti, M. B. & Gutierrez, C. *Arabidopsis* E2Fc functions in cell division and is degraded by the ubiquitin-SCF^{SKP2} pathway in response to light. *Plant Cell* **14**, 3057–3071 (2002).
- de Jager, S. M., Menges, M., Bauer, U. M. & Murray, J. A. H. *Arabidopsis* E2F1 binds a sequence present in the promoter of S-phase-regulated gene *AtCDC6* and is a member of a multigene family with differential activities. *Plant Mol. Biol.* **47**, 555–568 (2001).
- Mariconti, L. *et al.* The E2F family of transcription factors from *Arabidopsis thaliana*: novel and conserved components of the retinoblastoma/E2F pathway in plants. *J. Biol. Chem.* **277**, 9911–9919 (2002).
- Kosugi, S. & Ohashi, Y. Interaction of the *Arabidopsis* E2F and DP proteins confers their concomitant nuclear translocation and transactivation. *Plant Physiol.* **128**, 833–843 (2002).
- de Jager, S. *et al.* Dissecting regulatory pathways of G1/S control in *Arabidopsis*: common and distinct targets of CYCD3;1, E2Fa and E2Fc. *Plant Mol. Biol.* **71**, 345–365 (2009).
- Heckmann, S. *et al.* The E2F transcription factor family regulates CENH3 expression in *Arabidopsis thaliana*. *Plant J.* **68**, 646–656 (2011).
- del Pozo, J. C., Diaz-Trivino, S., Cisneros, N. & Gutierrez, C. The E2FC-DPB transcription factor controls cell division, endoreplication and lateral root formation in a SCFSKP2A-dependent manner. *Plant Signal. Behav.* **2**, 273–274 (2007).
- Yamaguchi, M. *et al.* VND-INTERACTING2, a NAC domain transcription factor, negatively regulates xylem vessel formation in *Arabidopsis*. *Plant Cell* **22**, 1249–1263 (2010).
- Yamaguchi, M. *et al.* VASCULAR-RELATED NAC-DOMAIN6 (VND6) and VND7 effectively induce transdifferentiation into xylem vessel elements under control of an induction system. *Plant Physiol.* **153**, 906–914 (2010).
- Wenkel, S., Emery, J., Hou, B.-H., Evans, M. M. S. & Barton, M. K. A feedback regulatory module formed by LITTLE ZIPPER and HD-ZIPIII genes. *Plant Cell* **19**, 3379–3390 (2007).
- Zhong, R. & Ye, Z.-H. MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol.* **53**, 368–380 (2012).
- Ohashi-Ito, K., Oda, Y. & Fukuda, H. *Arabidopsis* VASCULAR-RELATED NAC-DOMAIN6 directly regulates the genes that govern programmed cell death and secondary wall formation during xylem differentiation. *Plant Cell* **22**, 3461–3473 (2010).
- Pyo, H., Demura, T. & Fukuda, H. TERE, a novel *cis*-element responsible for a coordinated expression of genes related to programmed cell death and secondary wall formation during differentiation of tracheary elements. *Plant J.* **51**, 955–965 (2007).
- Dinneny, J. R. *et al.* Cell identity mediates the response of *Arabidopsis* roots to abiotic stress. *Science* **320**, 942–945 (2008).
- Iyer-Pascuzzi, A. S. *et al.* Cell identity regulators link development and stress responses in the *Arabidopsis* root. *Dev. Cell* **21**, 770–782 (2011).
- Gifford, M. L., Dean, A., Gutierrez, R. A., Coruzzi, G. M. & Birnbaum, K. D. Cell-specific nitrogen responses mediate developmental plasticity. *Proc. Natl Acad. Sci. USA* **105**, 803–808 (2008).
- Brady, S. M. *et al.* A stele-enriched gene regulatory network in the *Arabidopsis* root. *Mol. Syst. Biol.* **7**, 459 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. Tierney (University of Vermont) for 35S::GFP seeds, T. Demura for VND7 resources, M.K. Barton for REV:GR seeds, E.P. Spalding for advice on manuscript revision, and C. Gutierrez for E2Fc RNAi and E2Fc N-terminal deletion overexpressor seeds and useful discussion. This research was supported by the Office of Science (BER) Department of Energy Grant DE-FG02-08ER64700DE (to S.P.H. and S.A.K.), National Institute of General Medical Sciences of the National Institutes of Health under award numbers R01GM056006 and R02GM092412 (to S.A.K.), National Institute of Health (R01GM107311) and National Science Foundation (IOS-1036491 and IOS-1352478) to K.D., USDA CRIS 1907-21000-030 to D.W. and L.F., a Royal Society UK Fellowship (to S.E.A.), and UC Davis Startup Funds and a Hellman Fellowship (to S.M.B.).

Author Contributions M.T.-T., L.L. and M.d.L. contributed equally to this work. T.W.T. and A.G. contributed equally to this work. S.M.B. and S.P.H. contributed equally to this work. M.T.-T., L.L., M.d.L., S.M.B. and S.P.H. designed the research. M.T.-T., L.L., M.d.L., A.G., G.X., N.F.Y., G.M.T., M.T.V., R.L., P.P.H., C.W., and K.D. performed the research. M.T.-T., L.L., G.T., T.W.T., N.T., J.C., M.P., D.K., I.T., S.E.A., S.M.B. and S.P.H. analysed the data. L.Z., D.W., G.B., J.L.P.-P., and S.A.K. contributed new reagents/analytic tools. M.T.-T., L.L., G.M.T., S.M.B. and S.P.H. wrote the article. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.M.B. (sbrady@ucdavis.edu) and S.P.H. (hazen@bio.umass.edu).

METHODS

Yeast one-hybrid (Y1H) protein–DNA interaction assays. The root vascular-expressed transcription factor collection is described in ref. 7. The 1,663 transcription factor collection was assembled primarily from clones deposited in the *Arabidopsis* Biological Resource Center by various collaborative projects including the Peking-Yale Consortium³⁴, REGIA³⁵, TIGR³⁶, and the SSP Consortium³⁷. Translational fusions to the GAL4 activation domain were generated as described in ref. 38. A total of 1,663 *E. coli* strains harbouring different *Arabidopsis* transcription factors (Supplementary Table 5) were arrayed in 96-well plates and plasmids were prepared using the Promega Wizard SV 96 plasmid purification DNA system according to manufacturer's recommendations.

Root secondary cell wall gene promoters (2–3 kb of upstream regulatory region from the gene's translational start site, or the next gene, whichever comes first) were cloned and recombined with reporter genes according to ref. 33. Promoter sequences and primers used are described in Supplementary Table 1. AT1G30490, AT5G60690, AT2G34710, AT1G71930, AT1G62990 promoter sequences and primers are described in ref. 33, while the promoter sequences and primers for AT5G15630 are described in ref. 5. For dissection of cell wall biosynthesis promoters, approximately 1,000 bp of sequence upstream of the translational start site was tested for interactions with the transcription factor library. Three overlapping fragments of approximately equal and average size of 419 bp were independently cloned for each promoter according to ref. 38. The oligonucleotides used to amplify promoter fragments and details of their coordinates for *4CL1* (At1g51680), *CESA4/IRX5* (At5g44030), *CESA7/IRX3* (At5g17420), *CESA8/IRX1* (At4g18780), *COBL4/IRX6* (At5g15630), *HCT* (At5g48930), *IRX9* (At1g27600), *IRX14* (At4g36890), *KOR/IRX2* (At5g49720), *LAC4/IRX12* (At2g38080), and *REF8* (At2g40890) are described in Supplementary Table 6.

Root bait promoters were screened against the stele-expressed transcription factor collection using the Y1H protocol as previously described⁷. The 1,663 transcription factor library was transformed into each yeast strain and the β -galactosidase activity was determined as described in ref. 38, but in 384-well plates. Positive interactions were visually identified as incidence of yellow caused by the presence of ortho-nitrophenyl cleavage from colourless ortho-nitrophenyl- β -D-galactoside by β -galactosidase. The DNA bait strains were similarly tested for self-activation before screening by not transforming with prey vectors in the presence of thiamine. All interacting transcription factors were assembled into a cell wall interaction library and the screen was repeated to confirm the results and each clone was sequenced to reconfirm identity.

Statistical analysis for protein family enrichment. Enrichment was determined using the hypergeometric distribution online tool (<http://stattrek.com/>). The population size is the number of transcription factors in the xylem transcription factor collection while the successes within the population is the number of transcription factors within that transcription factor family in the xylem. The number of successes in the sample was the number of proteins belonging to that family, and the number in the sample is the total number of transcription factors within the network. The *A. thaliana* transcription factor list is as described in ref. 7.

Power graph compression approach. The power graph compression was performed using the algorithm as previously described¹⁴.

Plant material. The *E2Fc* RNAi line is described in ref. 23 and was verified by quantifying *E2Fc* transcript abundance relative to the Col-0 control using an *E2Fc* primer compared to an *ACTIN* control primer (Supplementary Table 1). *VND7::YFP* lines are described in ref. 39. The *VND7* glucocorticoid induction line is described in ref. 9. The *rev-5* loss-of-function mutant was described in ref. 40.

Cloning and insertion of the *4CL1* promoter into a pENTR p4-p1R donor vector was performed according to ref. 33 (for sequence, see Supplementary Table 1). The promoter was then recombined into binary vector pK7m24GW₃ along with pENTR 221 *ER-GFP::NOS*. The resulting *4CL1::GFP* vector was transformed into *Agrobacterium* strain GB3101. Col-0 plants were then transformed using the floral dip method.

Plant growth conditions. All plants were grown vertically on plates containing 1× Murashige and Skoog salt mixture, 1% sucrose, and 2.3 mM 2-(*N*-morpholino) ethanesulphonic acid (pH 5.8) in 1% agar. NaCl plates were made by adding 140 mM NaCl to this standard media. Iron control and deprivation media were made according to ref. 30. Plants grown on stress media (iron or salt) were first germinated on nylon mesh placed over control media for four days before transferring mesh with seedlings to iron deprivation or NaCl plates. Plants used for RNA isolation were also grown on nylon mesh placed over the agar to facilitate the collection of root material¹⁵.

Determination of crystalline cellulose. Roots of 7-day-old plants were harvested and lyophilized. Six to ten plates of seedlings grown at the same time on the same media were pooled to make a single biological replicate. Crystalline cellulose was measured according to ref. 41. After hydrolysis of non-cellulosic polysaccharides from an alcohol insoluble residue wall preparation with the Updegraff reagent

(acetic acid:nitric acids:water, 8:1:2 v/v), the remaining pellet was hydrolysed in 72% sulfuric acid. The resulting glucose quantity was determined by the anthrone method⁴².

Phloroglucinol staining. Five day after imbibition seedlings to be stained with phloroglucinol were fixed in a 3:1 95% ethanol:glacial acetic acid solution for 5 min. Samples were then transferred to a solution of 1% phloroglucinol in 50% HCl for 1–2 min. Whole seedlings were then mounted in 50% glycerol on slides and viewed using an Olympus Vanox microscope. Images were captured with a PIXERA Pro-600ES camera.

Confocal laser scanning microscopy. Confocal laser scanning microscopy was carried out on a Zeiss LSM700. Cell walls were stained using propidium iodide as previously described³⁰.

Transient protein–DNA interaction detection in tobacco. β -glucuronidase. For transient transactivation expression assays, the *VND7*, *GAL4*, and/or *CyclinB1* promoters were cloned into pGWB3 to generate GUS (β -glucuronidase gene) fusion reporters for *E2Fc* transcriptional activity. The *E2Fc* effector vector⁴³ (in PYL436) was provided by S. D. Kumar (UC Davis, CA). The effector and reporter constructs were transformed into *Agrobacterium tumefaciens* strain GV3101 and co-infiltrated with the p19 silencing inhibitor into 3-weeks-old *Nicotiana benthamiana* leaves at $A_{600\text{ nm}}$ 0.6:0:6:1, respectively. Leaves were harvested 3 days after agro-infiltration and homogenized in GUS extraction buffer (50 mM Na₂PO₄ pH 7, 10 mM Na₂EDTA, 0.1% SDS, 0.1% Triton TX-100 and 10 mM β -mercaptoethanol). Quantitative MUG fluorescent assay for GUS determination was performed using 100 μ g of protein/sample in 500 μ l of GUS assay buffer (1 mM 4-methyl umbelliferyl β -D-glucuronide, SIGMA, in Extraction Buffer). Samples were covered in aluminium foil and incubated at 37 °C. Reaction was stopped at different time points by transferring 50 μ l to a tube with 450 μ l of Stop Buffer (0.2 M Na₂CO₃). 4-methylumbelliferone fluorescence was determined using a Infinite 200 Pro-series reader (excitation at 365 nm, emission at 455 nm).

Luciferase (Fig. 2). Overnight cultures of *Agrobacterium* (GV3101, $D_{600\text{ nm}}$ = 0.6) carrying *VND7* promoter fused to *luciferase* (*LUC*) and 35S::*E2Fc* were prepared in infiltration medium (2 mM Na₃PO₄, 50 mM MES, 0.5% glucose, 100 μ M acetosyringone) at $D_{600\text{ nm}}$ = 0.1. Subsequently, cultures containing *VND7::LUC* and 35S::*E2Fc* at respective ratios of 1:0, 1:0.5, 1:1, 1:2, 1:5, and 1:10 were spot-infiltrated into 6–7-week-old *Nicotiana benthamiana* leaves. To prevent gene silencing, *Agrobacterium* strain carrying the pBIN19 suppressor from tomato bushy stunt virus was included in each of the combinations⁴⁴. The *LUC* activity was inspected at 72 to 96 h post infiltration using CCD camera (Andor Technology).

Luciferase imaging of *VND7::LUC* was performed as previously described with modifications⁴⁵. Briefly, tobacco leaves were cut off after 3 days of transient transformation and sprayed with 1 mM luciferin (Promega) in 0.01% Tween-80, then were imaged using an Andor DU434-BV CCD camera (Andor Technology). Images were acquired every 10 min for 12 pictures. Luciferase activity was quantified for a defined area as mean counts per pixel per exposure time using Andor Solis image analysis software (Andor Technology). Statistical analyses were performed using two-tailed Student's *t*-tests. The difference was considered significant if $P < 0.05$.

Luciferase (Fig. 4). A vector system was created to generate a single vector with the CaMV 35S constitutive promoter (35S) fused to a transcription factor, a promoter fragment fused to the firefly luciferase reporter gene, and 35S fused to the *Renilla* luciferase reporter gene. The constitutively expressed *Renilla* gene served as a control to normalize for transformation efficiency. This system includes one destination vector pLAH-LARm and three entry vectors pLAH-TF, pLAH-PROM and pLAH-VP6435T using MultiSite Gateway Pro Technology (Invitrogen) to simultaneously clone three DNA fragments (Extended Data Fig. 8). To develop the expression vector, promoter fragments and transcription factors were cloned, using the BP system (Invitrogen), into pDONR-P3-P2 and pDONR-P1-P4 to create pLAH-TF and pLAH-Prom, respectively. PacI-digested pMDC32 was ligated with the 2.427 kb pFLASH fragment following HindIII and SacI digestion to yield pLAR-L with the firefly luciferase (*LUC*) reporter gene. The 3 kb pRTL2-*Renilla* HindIII-digested fragment was inserted into SacI-digested pLAH-L to create pLAR-LR with both firefly *LUC* and *Renilla* luciferase (*REN*) genes. To generate pLAH-LAR, a SpeI-digested PCR fragment containing the *AmpR* gene amplified from pDEST22 was ligated with SpeI-digested pLAR-LR. To add the minimal CaMV 35S fragment (Mini35S) before the *LUC* reporter gene, the gateway cassette *ccdB/CmR* of pLAR-LAR was replaced by a HindIII-digested PCR fragment *Mini35S-ccdB-CmR* amplified from pMDC32 using specific primer pHindIII-Rv and primer Mini35S-attR2. The final destination vector is referred to as pLAH-LARm.

The protein coding regions of select transcription factor genes were amplified. Each amplified fragment was recombined with pDONR-P1-P4 vector by performing BP reactions to produce pLAH-TF. Target promoter fragments were amplified from *A. thaliana* genomic DNA using appropriate primers with attB3 and attB2 sites (Supplementary Table 10). Each amplified fragment was cloned into pDONR-P3-P2 vector by performing BP reactions to produce pLAH-PROM. A third pDONR

vector (pLAH-VP64Ter) was designed to create a carboxy-terminal fusion of the strong transcription activation domain VP64 to the transcription factor followed by the 35S transcription terminator (35St). A PCR fragments containing VP64 region and 35S terminator was amplified from pB7-VP64 using specific primers with attB4r and attB3r sites (Supplementary Table 10) into pDONR P4r-P3r to produce pLAH-VP6435T. Finally, the fully functional expression vector was generated by Gateway LR cloning of destination vector and the three entry clones: pLAH-LARM, pLAH-TF, and pLAH-VP64Ter (Extended Data Fig. 7).

Agrobacterium tumefaciens strain GV3103 (MP90) carrying expression constructs were grown in Luria-Bertani media with rifampicin and ampicillin and suspended in infiltration buffer 10 mM MES, pH 5.7, containing 10 mM MgCl₂ and 150 μM acetosyringone. The cultures were adjusted to a $D_{600\text{ nm}}$ of 0.8 and incubated at room temperature for at least 3 h before infiltration. The cultures were hand infiltrated using a 1 ml syringe into 3- to 4-week-old *N. benthamiana* leaves. Leaf samples were harvested 36 h after infiltration and assayed for luciferase activity according to manufacturer instructions using the Dual-Luciferase Reporter Assay Systems (Promega). Approximately 100 mg of tissue was frozen in liquid nitrogen and homogenized using a Retsch Mixer Mill MM400 for 1 min at 30 Hz. Ground tissue was then thawed in lysis buffer (0.1 M HEPES, pH 7.8, 1% Triton X-100, 1 mM CaCl₂ and 1 mM MgCl₂) at 25 °C for 15 min. Then 50 μl of Luciferase Assay Reagent II was added to 10 μl aliquots of the lysates to measure firefly luciferase activity, 1,000 ms integration time, using a Spectra Max M5/M5e plate reader to measure total light emission. Firefly luciferase activity was quenched with 50 μl of Stop & Glo Reagent, which contains *Renilla* luciferin substrate, also measured, 100 ms integration time, as total light emission. An expression vector containing part of the coding sequence (+X+Y) of the β-glucuronidase reporter gene rather than a transcription factor gene was used for baseline measurement of firefly luciferase activity. To estimate relative transcription factor affinity with each promoter fragment, three biological replicates of transcription factor expressing vectors were compared to the average results for the GUS expression vector. First, dividing firefly luciferase activity by *Renilla* luciferase activity normalized the transformation efficiency of each infiltrated leaf sample. Relative binding of the transcription factors to the promoter bait sequences was determined relative to the GUS control using a Student's *t*-test in R v2.11.0.

Electrophoretic mobility shift assays. To express recombinant NST2 or SND1 protein, coding sequence was cloned and fused to glutathione S-transferase tag in the pDONR211 vector and then transferred into pDEST15 (Invitrogen). *E. coli* strain BL21-AI (Invitrogen) transformed with pDEST15-GST:NST2 were grown in liquid media to a $D_{600\text{ nm}}$ of 0.4, treated with 0.2% L-arabinose to induce expression overnight and harvested by centrifugation the following day. Cells were treated with 1 mg ml⁻¹ lysozyme on ice for 30 min in minimal volume of 1× PBS buffer and lysed by sonication. Cell lysates were clarified by centrifugation and incubated with 100 μl of glutathione Sepharose beads (GE Healthcare) for 30 min at 4 °C with rotation. The beads were transferred to a column, washed with 10 volumes of 1× PBS. Protein was eluted in 100 mM Tris-HCl pH 8.0, 100 mM NaCl and 3 mg ml⁻¹ glutathione buffer and purified protein was resuspended in 50% glycerol and stored at -80 °C.

Three overlapping probes were generated for *CESA7*, *CESA8* and *KOR* promoters using the same oligonucleotides described in Supplementary Table 1, whereas three probes were generated for *CESA4* using the following primers: *CESA4*pr-1fwd, CACCGGGCCTTTGTGAAATTGATTTGGG; *CESA4*pr-1rev, TGTA TTTCTACTTTAGTCTTAC; *CESA4*pr-2fwd, CCAGATTTGGTAAAGTTTAT AAG; *CESA4*pr-2rev, GTGTCATAGAAAGCTTCAAG; *CESA4*pr-3fwd, TCTT ATGACACAAACCTTAGAC; *CESA4*pr-3rev, AACTGAGCTCTCGGAAGC AGAGCAG. Reactions were carried out in binding buffer (10 mM Tris, pH 7.5, 50 mM KCl, 1 mM DTT, 2.5% glycerol, 5 mM MgCl₂, 0.1% IGEPAL CA-630, and 0.05 μg μl⁻¹ calf thymus DNA). Following the addition of 150 ng of protein from the GST purification eluate, reactions were incubated at room temperature for 30 min. Protein-DNA complexes were separated from the free DNA on 1% agarose/1× TAE gels at 4 °C. The agarose gels were stained with ethidium bromide and bands visualized under ultraviolet light. For the titration of promoter DNA with NST2 protein, *CESA4* promoter fragment-2 DNA and *KOR* promoter fragment-1 DNA in 30 ng were titrated with increasing amounts of NST2 protein: 25, 50, 150, 300, and 600 ng. Binding reaction and the separation of protein-DNA complexes were carried out as described above.

Chromatin immunoprecipitation of NST2. Chromatin immunoprecipitation was conducted as described in ref. 46 with the following modifications. Roughly 5 g (fresh weight) whole stems from six-week-old *Arabidopsis* were harvested and crosslinked for 15 min under vacuum in crosslinking buffer (10 mM Tris, pH 8.0, 1 mM EDTA, 250 mM sucrose, 1 mM PMSF and 1% formaldehyde). Technical replicates containing approximately 1.5 mg DNA were resuspended in 800 μl SII buffer, incubated with 2 μg anti-GFP antibody (ab290, Abcam) bound to Protein G Dynabeads (Invitrogen) for 1.5 h at 4 °C and then washed five times with SII

buffer. Chromatin was eluted from the beads twice at 65 °C with Stop buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 20 mM EDTA and 1% SDS). RNase- and DNase-free glycogen (2 μg) (Boehringer Mannheim) was added to the input and eluted chromatin before they were incubated with DNase- and RNase-free proteinase K (Invitrogen) at 65 °C overnight and then treated with 2 μg RNase A (Qiagen) for 1 h at 37 °C. DNA was purified by using Qiagen PCR Purification kit and resuspended in 100 μl H₂O. Quantitative PCR reactions of the technical replicates were performed using Quantifast SYBR Green PCR Kit (Qiagen), with the following PCR conditions: 2 min at 95 °C, followed by 40 cycles of 15 s at 95 °C, 15 s at 55 °C and 20 s at 68 °C. Primers used in this study are listed in Supplementary Table 4. Results were normalized to the input DNA, using the following equation: $100 \times 2^{(\text{Ct input} - 3.32 - \text{Ct ChIP})}$.

Quantitative RT-PCR. Primers for qRT-PCR were designed to amplify a 100 bp region (or a 400 bp region for *REV*, *PHB*, and *PHV* transcripts due to sequence similarity) on the 3' end of each transcript³³. Primer sets used for qRT-PCR are listed in Supplementary Table 1. Each plate was considered a biological replicate and Columbian and reference genotypes were plated on the same plate. Five days after imbibition, total RNA was extracted from seedling roots using an RNeasy Kit (QIAGEN). cDNA was synthesized by treatment with reverse transcriptase and oligo(dT) primer (SuperScript III First-Strand Synthesis System; Invitrogen). qRT-PCR was performed in an iCycler iQ Real-Time PCR Detection System (Bio-Rad) using the Bio-rad iQ SYBR green Supermix. Gene expression was measured between wild-type and mutant pairs across at least two biological replicates with three technical replicates using the $\Delta - \Delta C_T$ method³⁰.

VND7 induction experiments. VND7-VP16-GR⁹ plants were grown vertically on sterile mesh placed on top of MS media with sucrose. Five days after imbibition, seedlings were transferred, with the mesh, to MS media containing 10 μM dexamethasone and roots were collected for qRT-PCR (RNeasy Kit; Qiagen) after 0, 1, 2, 3, or 4 h on dexamethasone ($n = 3$). As a positive control, upregulation of *MYB46* expression was confirmed using qRT-PCR.

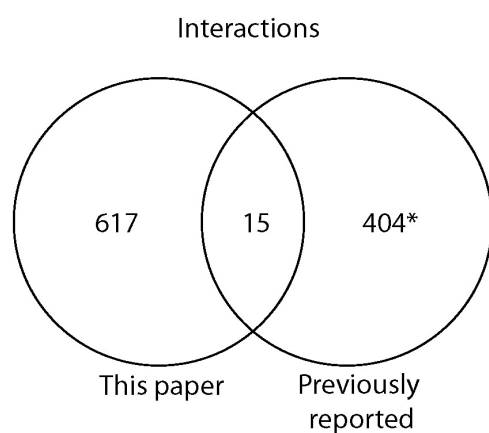
Nitrogen influx, salt stress, iron deprivation, sulphur stress, pH stress analysis. The data sets used contained mean expression values for each gene in both control and treatment, and a *q* value for each gene indicating the significance of the hypothesis that the expression values of control and treatment are drawn from distributions with the same means. These data sets were filtered to extract only those genes whose *q* value was ≤ 0.01 and whose fold change between mean expression values was ≥ 1.5 in either direction. Fisher's exact test was used to test whether the number of such genes is overrepresented in the xylem cell specification and differentiation gene regulatory network.

Gene regulatory network inference. Expression data³⁰ were used, after normalization with the mmgMOS method from the PUMA R package⁴⁷. The supervised regulatory interactions network was constructed using SIRENE⁴⁸. The directionality of the interactions is defined by the protein-DNA interactions from Y1H data. The interaction sign (activation or repression) is derived by Pearson's correlation coefficient for each protein-DNA interaction. The analysis performed was categorized as (1) supervised tier Ia, network inferred with SIRENE with the provided Y1H gene regulatory connections and the corresponding gene expression profiles (16 genes, 4 transcription factors); (2) supervised tier Ib, an additional three verified connections from the supervised tier Ia and unsupervised tier I were considered in the inference. The unsupervised regulatory interaction network was constructed using the consensus from four different gene regulatory network inference methods, GENIE3⁴⁹, Inferelator⁵⁰, TIGRESS⁵¹ and ANOverence⁵². The data used were the same as the supervised TIERIa network. The default parameters were used in all methods and a rank-based method was used to build the consensus network as in ref. 53.

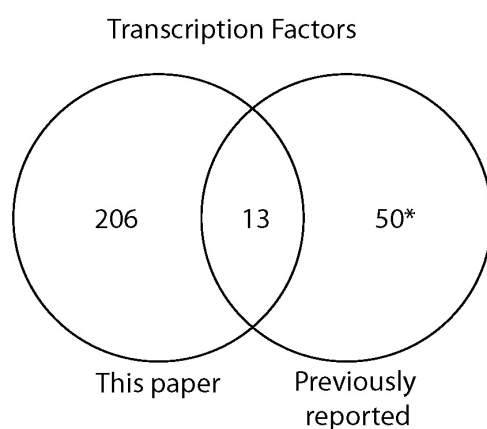
34. Gong, W. *et al.* Genome-wide ORFeome cloning and analysis of *Arabidopsis* transcription factor genes. *Plant Physiol.* **135**, 773–782 (2004).
35. Paz-Ares, J. & The REGIA Consortium. REGIA, an EU project on functional genomics of transcription factors from *Arabidopsis thaliana*. *Comp. Funct. Genomics* **3**, 102–108 (2002).
36. Underwood, B. A., Vanderhaeghen, R., Whitford, R., Town, C. D. & Hilson, P. Simultaneous high-throughput recombinational cloning of open reading frames in closed and open configurations. *Plant Biotechnol. J.* **4**, 317–324 (2006).
37. Yamada, K., Lim, J., Dale, J. & Chen, H. Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* **302**, 842–846 (2003).
38. Prunedu-Paz, J. L., Breton, G., Para, A. & Kay, S. A. A functional genomics approach reveals CHE as a component of the *Arabidopsis* circadian clock. *Science* **323**, 1481–1485 (2009).
39. Kubo, M. *et al.* Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.* **19**, 1855–1860 (2005).
40. Hawker, N. P. & Bowman, J. L. Roles for class III HD-Zip and KANADI genes in *Arabidopsis* root development. *Plant Physiol.* **135**, 2261–2270 (2004).
41. Updegraff, D. Semimicro determination of cellulose in biological materials. *Anal. Biochem.* **32**, 420–424 (1969).

42. Scott, T. A. & Melvin, E. H. Determination of dextran with anthrone. *Anal. Chem.* **25**, 1656–1661 (1953).
43. Liu, Y., Burch-Smith, T., Schiff, M., Feng, S. & Dinesh-Kumar, S. P. Molecular chaperone Hsp90 associates with resistance protein N and its signaling proteins SGT1 and Rar1 to modulate an innate immune response in plants. *J. Biol. Chem.* **279**, 2101–2108 (2004).
44. Voinnet, O., Rivas, S., Mestre, P. & Baulcombe, D. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J.* **33**, 949–956 (2003).
45. Walley, J. W. *et al.* Mechanical stress induces biotic and abiotic stress responses via a novel *cis*-element. *PLoS Genet.* **3**, e172 (2007).
46. Nusinow, D. A. *et al.* The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature* **475**, 398–402 (2011).
47. Pearson, R. D. *et al.* puma: a Bioconductor package for propagating uncertainty in microarray analysis. *BMC Bioinformatics* **10**, 211 (2009).
48. Mordelet, F. & Vert, J.-P. SIRENE: supervised inference of regulatory networks. *Bioinformatics* **24**, i76–i82 (2008).
49. Huynh-Thu, V. A., Irrthum, A., Wehenkel, L. & Geurts, P. Inferring regulatory networks from expression data using tree-based methods. *PLoS ONE* **5**, e12776 (2010).
50. Greenfield, A., Madar, A., Ostrer, H. & Bonneau, R. DREAM4: Combining genetic and dynamic information to identify biological networks and dynamical models. *PLoS ONE* **5**, e13397 (2010).
51. Haury, A.-C., Mordelet, F., Vera-Licona, P. & Vert, J.-P. TIGRESS: Trustful Inference of Gene REgulation using Stability Selection. *BMC Syst. Biol.* **6**, 145 (2012).
52. Küffner, R., Petri, T., Tavakkolkhah, P., Windhager, L. & Zimmer, R. Inferring gene regulatory networks by ANOVA. *Bioinformatics* **28**, 1376–1382 (2012).
53. Marbach, D. *et al.* Wisdom of crowds for robust gene network inference. *Nature Methods* **9**, 796–804 (2012).

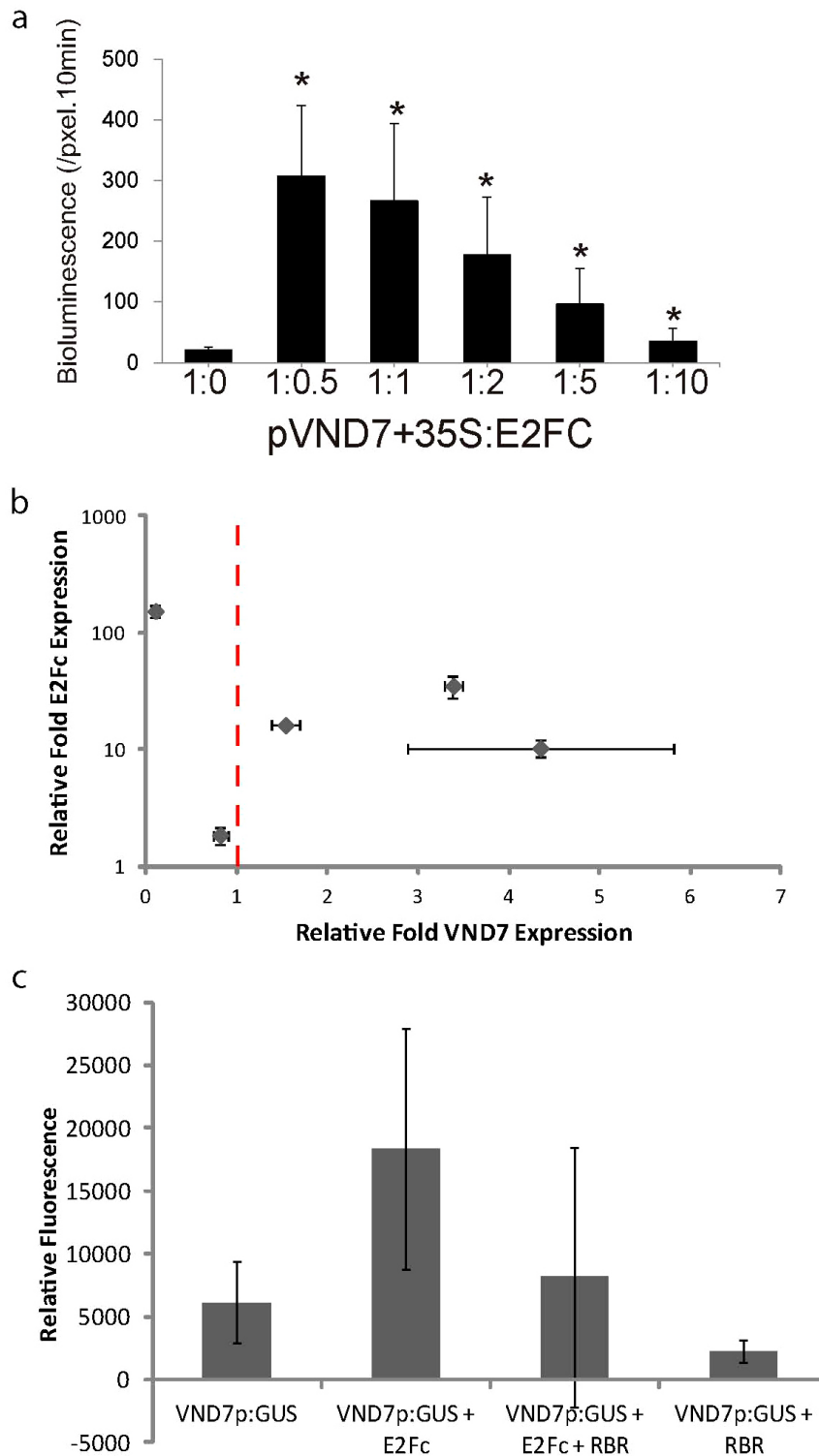
a



b

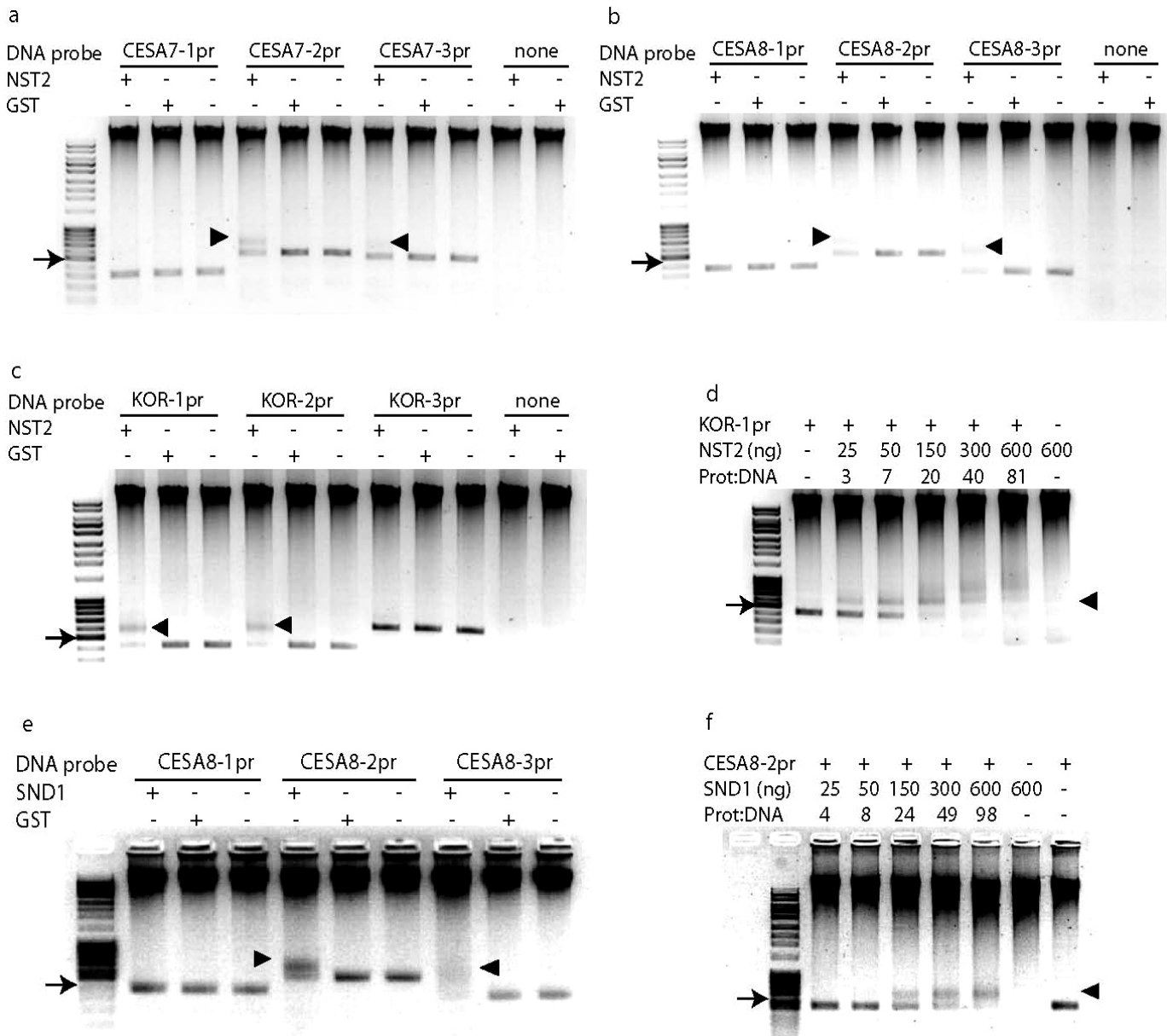


Extended Data Figure 1 | Number of novel and previously described protein–DNA interactions and transcription factors involved in secondary cell wall biosynthesis and xylem development. a, b, Venn diagrams of overlap between previously reported¹⁹ interactions (a) or transcription factors (b) and those of the xylem-specific gene regulatory network. *Includes genes that were not included in the yeast one hybrid screen.



Extended Data Figure 2 | Activation or repression of *VND7* by E2Fc is dynamic and dose-dependent. **a**, Intensity of LUC bioluminescence quantified using Andor Solis image analysis software. Data are means \pm s.d. ($n = 20$). Asterisks denote significance at $P < 0.05$ determined by Student's t -test. **b**, Quantitative PCR with reverse transcription of *E2Fc* and *VND7* transcripts in ΔN -E2Fc (*E2Fc* overexpressor line lacking the N-terminal domain) expressing plants versus Col-0 control. Red dashed line marks the

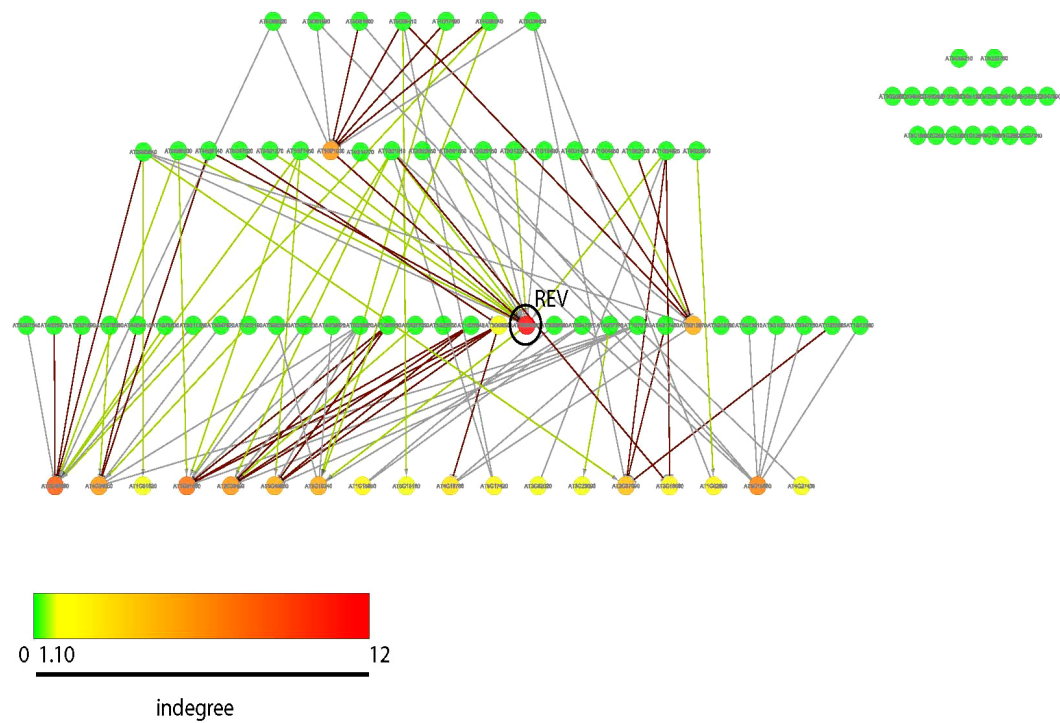
point at which *VND7* is unchanged compared to control. Each data point is an individual biological replicate with 3 technical replicates. **c**, 3-week-old tobacco leaves were infiltrated with the p19 silencing inhibitor and either the reporter *VND7p::GUS* or *VND7p::GUS* and either *35S::E2Fc::MYC* or *35S::RBR::GFP*, or both. Extracted protein was then used in a quantitative MUG fluorescent assay, where relative fluorescence was measured 60 min after incubation with substrate. Data are means \pm s.d., $n = 3$.



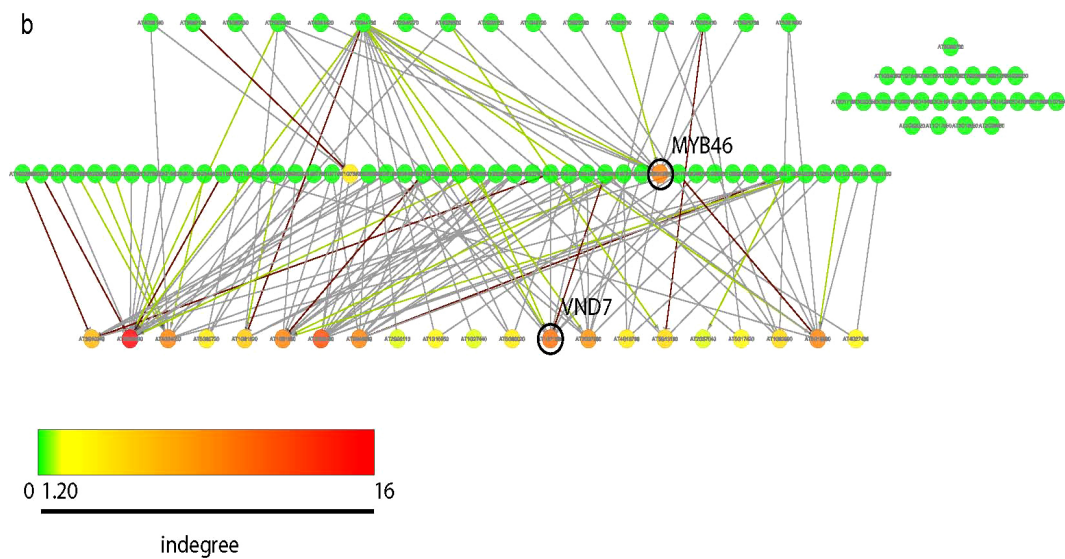
Extended Data Figure 3 | Binding of NST2 and SND1 to fragments of *CESA7*, *CESA8*, and *KOR* promoters. a–f, Electrophoretic mobility shift assays showing NST2 (a–d) and SND1 (e–f) protein specifically binds the

promoters of cellulose-associated genes. Probe was incubated in the absence or presence of GST or GST:SND1 protein extracts. The arrowheads indicate the specific protein–DNA complexes, while arrows indicate free probe.

a

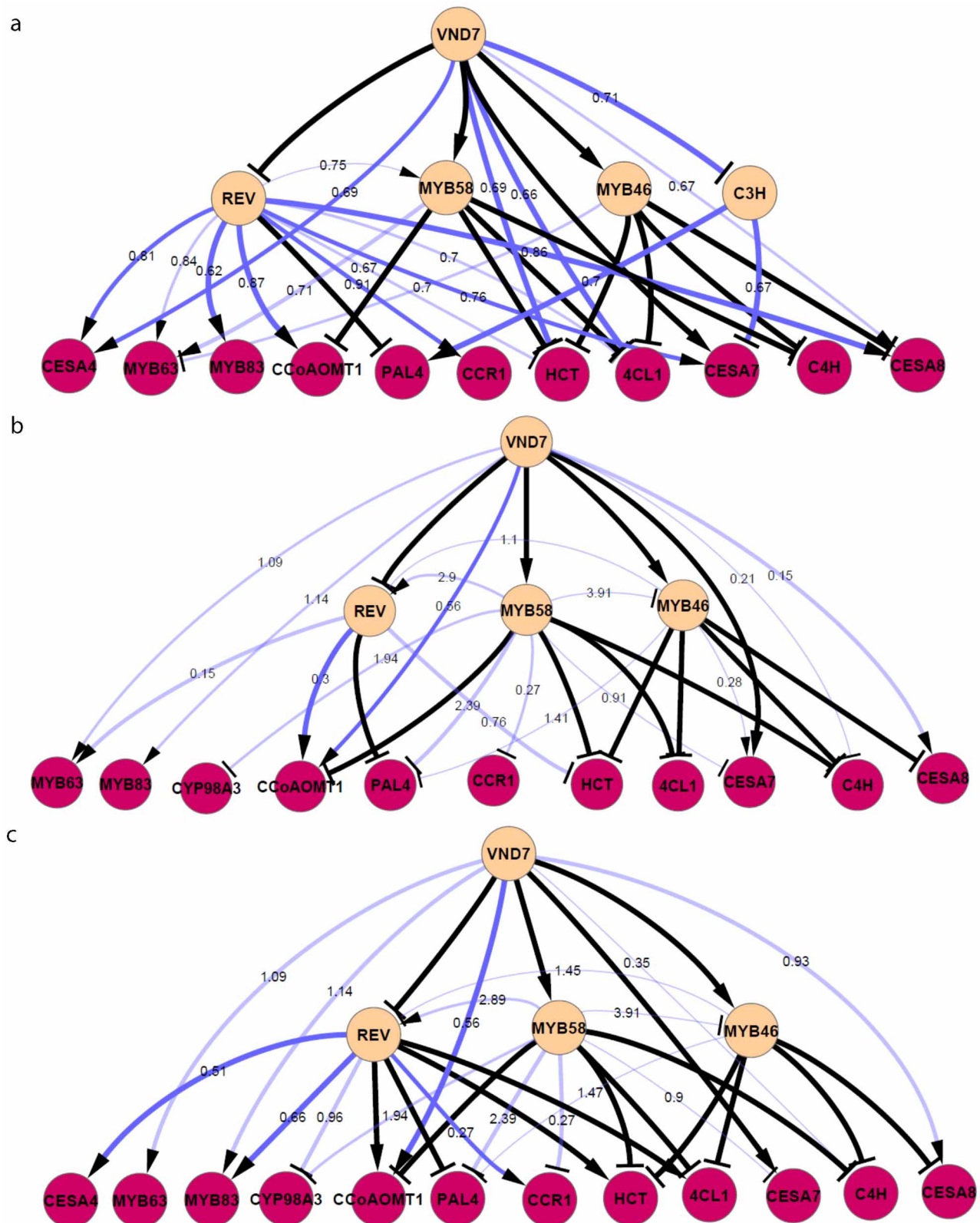


b



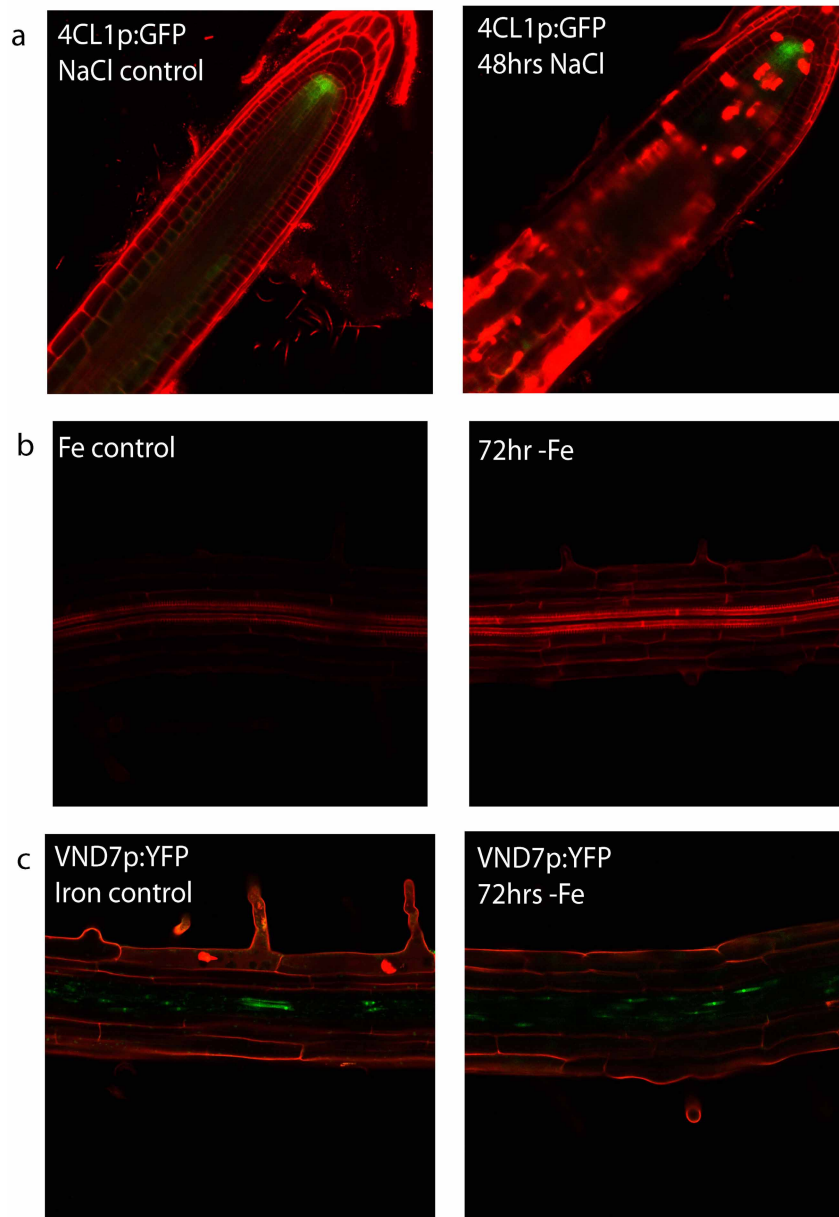
Extended Data Figure 4 | Sub-networks of network genes differentially expressed in response to iron deprivation of high salinity. a, b, Sub-network of genes with q values of ≤ 0.01 and whose fold change between mean expression values was ≥ 1.5 in either direction in iron deprivation (a) or high

NaCl (b) stress microarray data set. Nodes are coloured according to in-degree as shown on scale bars below sub-networks. Transcription factors with the highest in-degree are labelled and indicated with a black circle.



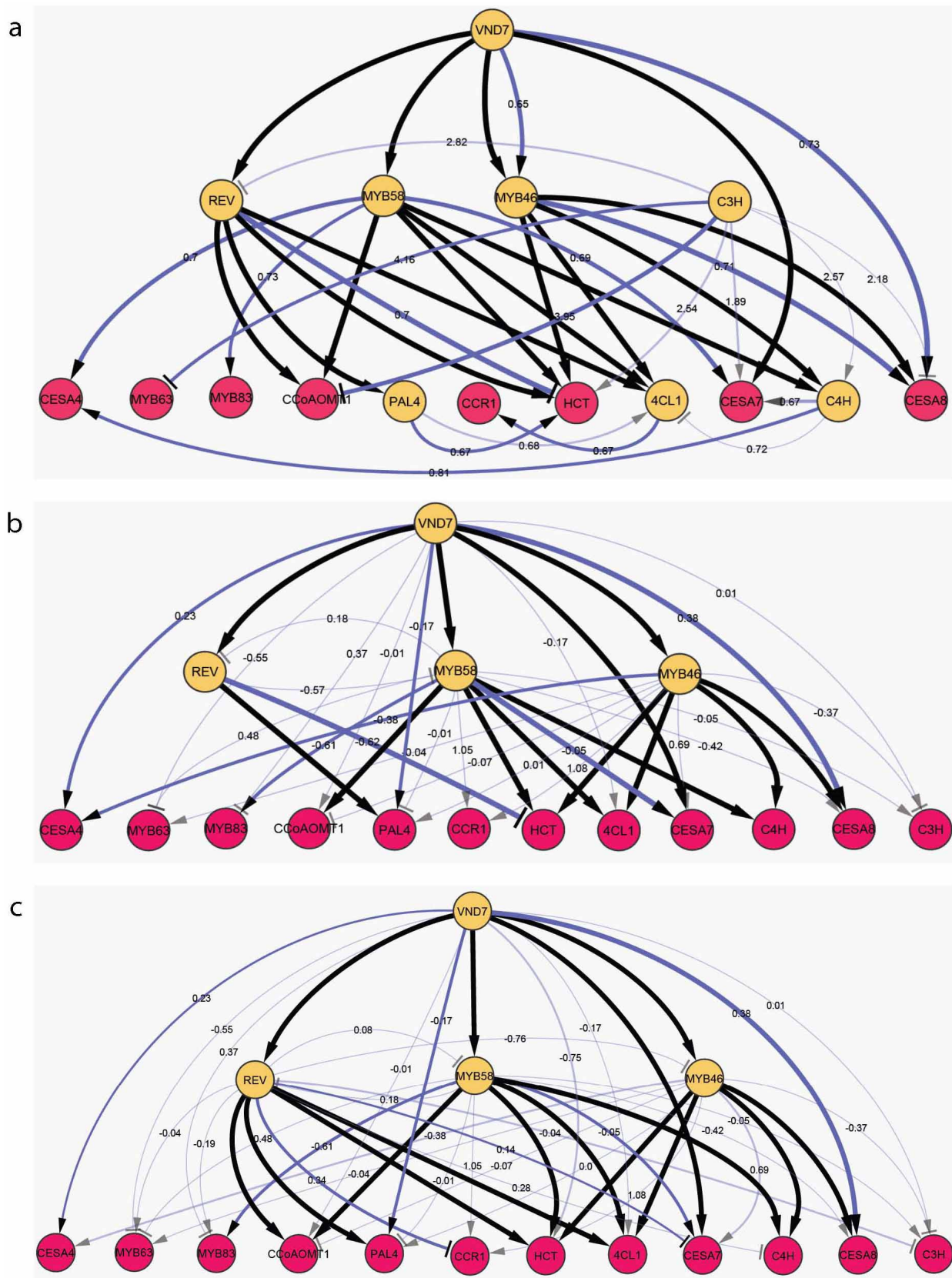
Extended Data Figure 5 | The reconstructed gene regulatory consensus network based on analysis of the iron-deprivation expression data set by different network inference methods. a, Unsupervised; b, supervised in the first pass; c, supervised after the validated two connections have been added in the training set. Edge transparency denotes $P \leq 0.06$ for the Pearson

correlation coefficient (PCC); edge width is proportional to PCC; edge value corresponds to the total edge score; a greater value corresponds to a more significant score. Yellow and red nodes correspond to transcription factor and target gene nodes, respectively; black and blue edges denote Y1H-derived and inferred interactions, respectively.



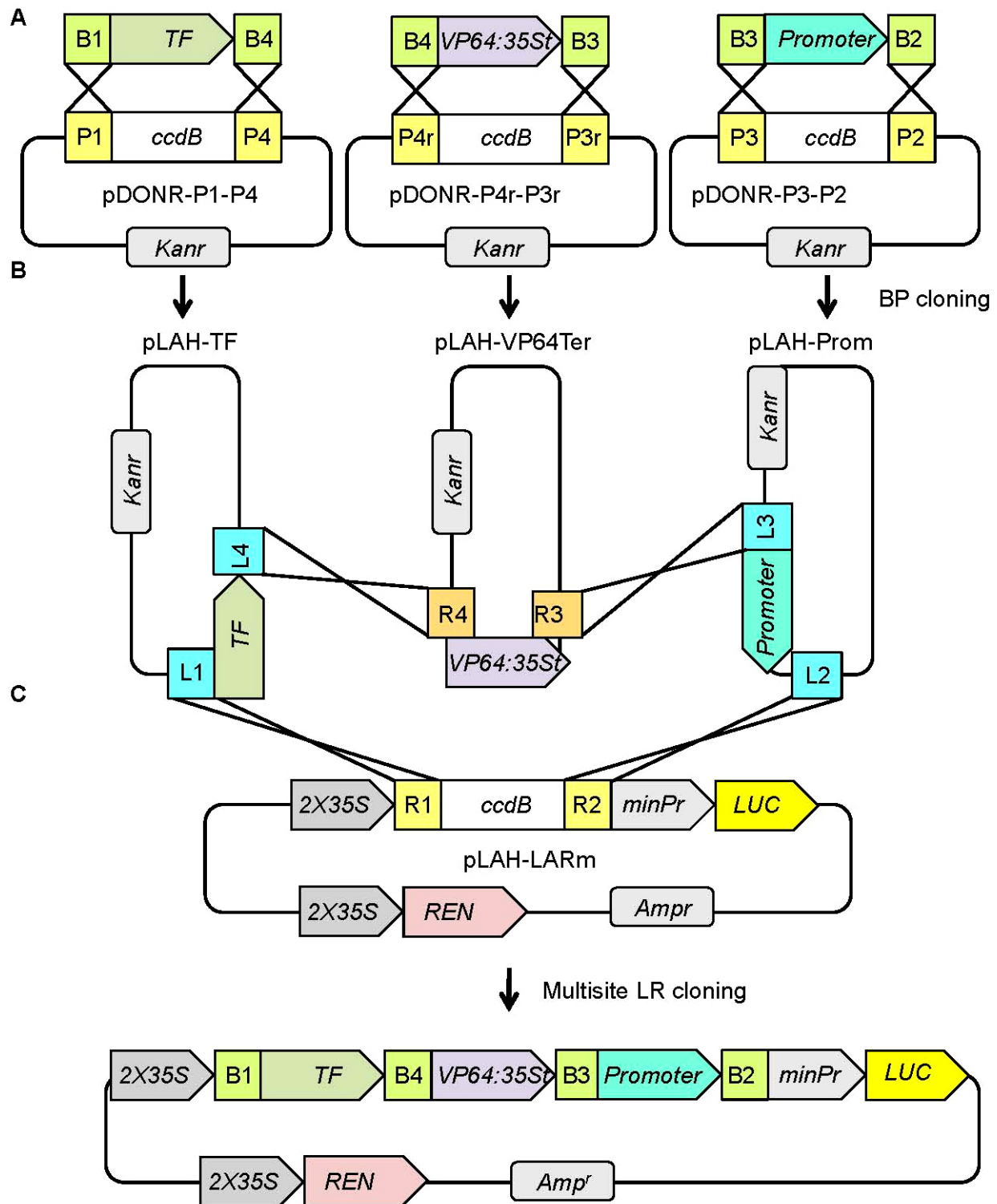
Extended Data Figure 6 | Iron deprivation and NaCl stress influences lignin and phenylpropanoid biosynthesis associated gene expression. **a**, No change was observed in the expression of *4CL1::GFP* in 4 days after imbibition (DAI) roots transferred to a control media (left, $n = 4$) or media with 140 mM NaCl for 48 h (right, $n = 4$). **b**, Increased fuchsin staining of xylem cells as well

as of cell walls of non-vascular cells in 4 DAI roots transferred to a control media (left) or media with an iron chelator for 72 h (right). **c**, No change was observed in the expression of *VND7::YFP* in 4 DAI roots transferred to a control media (left, $n = 4$) or media with an iron chelator for 72 h (right, $n = 5$).



Extended Data Figure 7 | The reconstructed gene regulatory consensus network based on analysis of the salt-stress expression data set by different network inference methods. a, Unsupervised; b, supervised in the first pass; c, supervised after the validated two connections have been added in the training set. Edge transparency denotes $P \leq 0.06$ for the Pearson correlation

coefficient (PCC); edge width is proportional to PCC; edge value corresponds to the total edge score; a greater value corresponds to a more significant score. Yellow and red nodes correspond to transcription factor and target gene nodes, respectively; black and blue edges denote YIH-derived and inferred interactions, respectively.



Extended Data Figure 8 | Schematic diagram of dual-luciferase reporter vector development. **a**, Three distinct donor vectors harbouring either the transcription factor, VP64 activation domain fused to the 35S minimal

promoter, or a promoter fragment. **b**, The dual reporter vector, pLAH-LARm, is then recombined with the three donor vectors to generate the single reporter vector (**c**).

Comprehensive genomic characterization of head and neck squamous cell carcinomas

The Cancer Genome Atlas Network*

The Cancer Genome Atlas profiled 279 head and neck squamous cell carcinomas (HNSCCs) to provide a comprehensive landscape of somatic genomic alterations. Here we show that human-papillomavirus-associated tumours are dominated by helical domain mutations of the oncogene *PIK3CA*, novel alterations involving loss of *TRAF3*, and amplification of the cell cycle gene *E2F1*. Smoking-related HNSCCs demonstrate near universal loss-of-function *TP53* mutations and *CDKN2A* inactivation with frequent copy number alterations including amplification of 3q26/28 and 11q13/22. A subgroup of oral cavity tumours with favourable clinical outcomes displayed infrequent copy number alterations in conjunction with activating mutations of *HRAS* or *PIK3CA*, coupled with inactivating mutations of *CASP8*, *NOTCH1* and *TP53*. Other distinct subgroups contained loss-of-function alterations of the chromatin modifier *NSD1*, *WNT* pathway genes *AJUBA* and *FAT1*, and activation of oxidative stress factor *NFE2L2*, mainly in laryngeal tumours. Therapeutic candidate alterations were identified in most HNSCCs.

HNSCCs affect ~600,000 patients per year worldwide¹. They are characterized by phenotypic, aetiological, biological and clinical heterogeneity. Smoking is implicated in the rise of HNSCC in developing countries, and the role of human papillomavirus (HPV) is emerging as an important factor in the rise of oropharyngeal tumours affecting non-smokers in developed countries². Despite surgery, radiation and chemotherapy, approximately half of all patients will die of the disease. Risk stratification for HNSCC is by anatomic site, stage and histological characteristics of the tumour. Except for HPV status, numerous molecular and clinical risk factors that have been investigated have limited clinical utility.

Published genome-wide profiling studies of HNSCC^{3,4} are limited to single platforms. To generate an integrated genomic annotation of molecular alterations in HNSCC, The Cancer Genome Atlas (TCGA) has undertaken a comprehensive multi-platform characterization of 500 tumours with the a priori hypothesis of detecting somatic variants present in at least 5% of samples. Here, we report the results for analyses from the first 279 patients with complete data.

Samples and clinical data

The cohort consists primarily of tumours from the oral cavity ($n = 172$ out of 279, 62%), oropharynx ($n = 33$ out of 279, 12%), and laryngeal sites ($n = 72$ out of 279, 26%) (Supplementary Information section 1, Supplementary Table 1.1 and Supplementary Data 1.1). Most patients were male ($n = 203$ out of 279, 73%) and heavy smokers (mean pack years = 51). Samples were classified as HPV-positive using an empiric definition of $>1,000$ mapped RNA sequencing (RNA-Seq) reads, primarily aligning to viral genes *E6* and *E7* (Supplementary Information section 1.2 and Supplementary Fig. 1.1). The HPV status by mapping of RNA-Seq reads was concordant with the genomic, sequencing and molecular data, and indicated that 36 tumours were HPV(+) and 243 were HPV(−) (Supplementary Information section 1.2, Supplementary Fig. 1.1 and Supplementary Data 1.2). Of 33 oropharyngeal tumours, 64% were positive for HPV, compared to 6% of 246 non-oropharyngeal tumours. Molecular HPV signatures were identified using microRNA

(miRNA), DNA methylation, gene expression and somatic nucleotide substitutions (Supplementary Information section 1.2 and Supplementary Figs 1.1–1.3). HPV(+) tumours exhibited infrequent mutations in *TP53* or genetic alterations in *CDKN2A*. We evaluated outcome by site, stage, HPV status, molecular subtypes and putative biomarkers (Supplementary Information section 1.3 and Supplementary Figs 1.4 and 1.5). HPV(+) and interestingly patients with HPV(−), *TP53* wild-type tumours demonstrated favourable outcomes compared to *TP53* mutants and 11q13/*CCND1*-amplified tumours.

DNA and RNA structural alterations

Most tumours demonstrated copy number alterations (CNAs) including losses of 3p and 8p, and gains of 3q, 5p and 8q chromosomal regions (Fig. 1a, Supplementary Fig. 2.1 and Supplementary Information section 2) resembling lung squamous cell carcinomas (LUSCs)⁵ (Fig. 1a and Supplementary Figs 2.1 and 2.2). HNSCC genomes showed high instability with a mean of 141 CNAs (amplifications or deletions) from microarray data and 62 structural aberrations (chromosomal fusions) per tumour by 'high coverage' whole-genome sequencing ($n = 29$) (Supplementary Information section 2.2). We observed 39 regions of recurrent copy number loss and 23 regions of recurrent copy number gain ($q < 0.1$, Supplementary Data 2.1 and 2.2). Both HPV(+) and (−) tumours contained recurrent focal amplifications for 3q26/28, a region involving squamous lineage transcription factors *TP63* and *SOX2* and the oncogene *PIK3CA* (Fig. 1b and Supplementary Fig. 2.3).

HPV(+) tumours were distinguished by novel recurrent deletions ($n = 5$ out of 36, 14%) and truncating mutations ($n = 3$ out of 36, 8%) of TNF receptor-associated factor 3 (*TRAF3*) (Supplementary Figs 2.3 and 2.4, and Supplementary Data 2.1). *TRAF3* is implicated in innate and acquired anti-viral responses⁶ including Epstein–Barr, HPV and human immunodeficiency virus (HIV)^{7–9}, while loss promotes aberrant NF- κ B signalling¹⁰. Although *TRAF3* inactivation has been reported in haematological malignancies and nasopharyngeal carcinoma^{11,12}, to our knowledge this is the first evidence linking *TRAF3* to HPV-associated

*Lists of participants and their affiliations appear at the end of the paper.

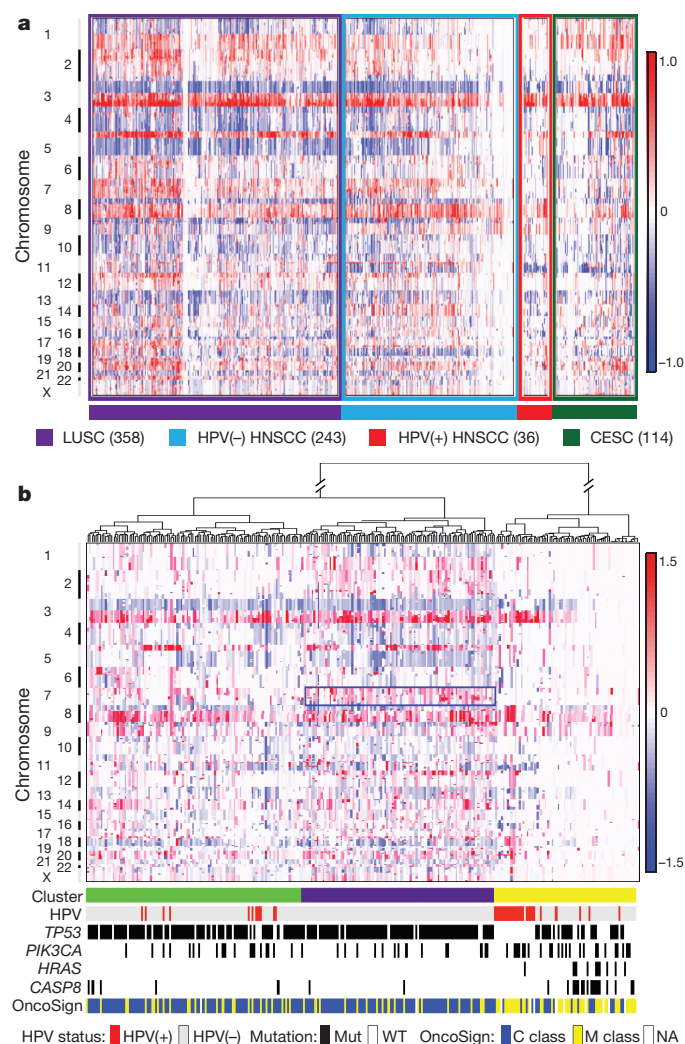


Figure 1 | DNA copy number alterations. **a**, Copy number alterations by anatomic site and HPV status for squamous cancers. Lung squamous cell carcinoma (LUSC, $n = 358$) and cervical squamous cell carcinoma (CESC, $n = 114$). **b**, Unsupervised analysis of copy number alteration of HNSCC ($n = 279$) with associated characteristics. The rectangle indicates chromosome 7 amplifications in the purple cluster. NA, not available.

carcinomas. HPV(+) tumours were also notable for focal amplification of *E2F1* and an intact 9p21.3 region containing the *CDKN2A* gene commonly deleted in HPV(-) tumours.

HPV(-) tumours featured novel co-amplifications of 11q13 (*CCND1*, *FADD* and *CTTN*) and 11q22 (*BIRC2* and *YAP1*), which also contain genes implicated in cell death/NF- κ B and Hippo pathways. HPV(-) tumours featured novel focal deletions in the nuclear set domain gene (*NSD1*) and tumour suppressor genes (for example, *FAT1*, *NOTCH1*, *SMAD4* and *CDKN2A*; Supplementary Fig. 2.3). Recurrent focal amplifications in receptor tyrosine kinases (for example, *EGFR*, *ERBB2* and *FGFR1*) also predominated in HPV(-) tumours. Notably, unsupervised clustering analysis of CNAs identified a mutually exclusive subset of predominantly oral cavity tumours with reduced CNAs, a pattern recently described in cancer as 'M' class (tumours driven by mutation rather than CNA)¹³ (Fig. 1b). This subset in particular contained a new three-gene pattern of activating mutations in *HRAS*, frequently with inactivating *CASP8* mutations, and wild-type *TP53*. We confirmed a previously reported favourable clinical outcome in tumours with few CNAs¹⁴. The three-gene constellation of wild-type *TP53* with mutant *HRAS* and *CASP8* suggested an alternative tumorigenesis pathway involving RAS and/or alterations in cell death/NF- κ B¹⁵. Unsupervised

analysis also suggested that clustering was a function of chromosome 7 amplification (including the *EGFR* locus) in a manner that largely excluded HPV(+) tumours.

To detect additional structural alterations, we interrogated whole-genome and RNA-Seq data (Supplementary Information section 3, Supplementary Data 3.1 and Supplementary Fig. 3.1). Known fusion oncogenes reported in solid tumours including those involving the *ALK*, *ROS* or *RET* genes were not observed in HNSCC. Previously reported *FGFR3-TACC3* fusions were present in two HPV(+) tumours (Supplementary Fig. 3.2). Only 1 out of 279 patients showed evidence of the type III isoform of *EGFR* (vIII), previously described in HNSCC¹⁶ (Supplementary Fig. 3.3). Although our investigation did not identify additional novel oncogenic fusions, several tumours demonstrated exon 1 of *EGFR* or *FGFR3* fused to non-recurrent partners, suggesting potential promoter swaps for the partner genes (Supplementary Data 3.1). A low prevalence of an alternative *MET* transcript with skipped exon 14 was identified in two HPV(-) tumours (Supplementary Fig. 3.4); this finding was reported to be an activating event in non-small cell lung cancer¹⁷. Structural alterations (homozygous deletions, intra- and inter-chromosomal fusions) were more commonly associated with loss of function in tumour suppressor genes, most prominently *CDKN2A* (Supplementary Figs 3.5 and 3.6), followed by *TP53*, *RBI*, *NOTCH1* and *FAT1* (Supplementary Figs 3.7–3.9), than with protein-coding fusion events. RNA-Seq data (Supplementary Data 3.3) demonstrated evidence of alternative splicing in genes not previously described in HNSCC including kallikrein 12 (*KLK12*) (Supplementary Fig. 3.11), as well as genes such as *TP63* with known importance in HNSCC (Supplementary Fig. 3.12).

By DNA analysis, most HPV(+) tumours demonstrated clear evidence of host genome integration, usually in a single genomic location per sample and almost always in association with amplifications of the host genome (Supplementary Fig. 3.10 and Supplementary Data 3.2). Interrogation of RNA transcripts confirmed transcription across the viral-human integration locus. However, none of the genes involved were recurrent, suggesting no single driver mechanism related to HPV integration. Similarly, none of the integration sites involved the *MYC* gene as reported in HPV(+) cell lines¹⁸.

Somatic mutations

Whole-exome sequencing identified somatically mutated genes, many located in regions of CNAs and annotated in the COSMIC database¹⁹ (Fig. 2). The mean sequencing coverage across targeted bases was 95 \times , with 82% of target bases above 30 \times coverage. In 279 samples, 12,159 synonymous somatic variants, 37,061 non-synonymous somatic variants, and 2,579 germline single base substitutions from the single nucleotide polymorphism database (dbSNP)²⁰ were detected (Supplementary Information section 4). Targeted re-sequencing of 394 unique regions (Supplementary Fig. 4.1) validated 99% of mutations. Interrogation of RNA for expression of the mutated alleles confirmed the variant in 86% of cases (Supplementary Information section 3.2 and Supplementary Fig. 3.1). In contrast to previous reports, the mutation rates did not differ by HPV status, although transversions at CpG sites were more frequent in HPV(-) tumours and a predominance of TpC mutations were noted in HPV(+) cases³ (Supplementary Fig. 1.1). Mutations were statistically enriched in 11 genes (Fig. 2). Among inactivating mutations (premature termination of the protein by nonsense, frameshift or splice-site mutations), four genes segregated exclusively or predominantly in HPV(-) tumours. Two were associated with cell cycle and survival (*CDKN2A* ($P < 0.01$) and *TP53* ($P < 0.01$)) and two were linked to Wnt/ β -catenin signalling (*FAT1* ($P < 0.01$) and *AJUBA* ($P = 0.14$))^{21,22}. We observed *TP53* mutation among HPV(-) samples at higher rates (86%) than have been previously reported¹⁹, while only 1 out of 36 HPV(+) cases had a non-synonymous *TP53* mutation. Previously unreported somatic mutations and deletions of *AJUBA* were primarily 5' inactivating events and clustered missense mutations in the functional LIM domain (Supplementary Fig. 4.2). *AJUBA* is a

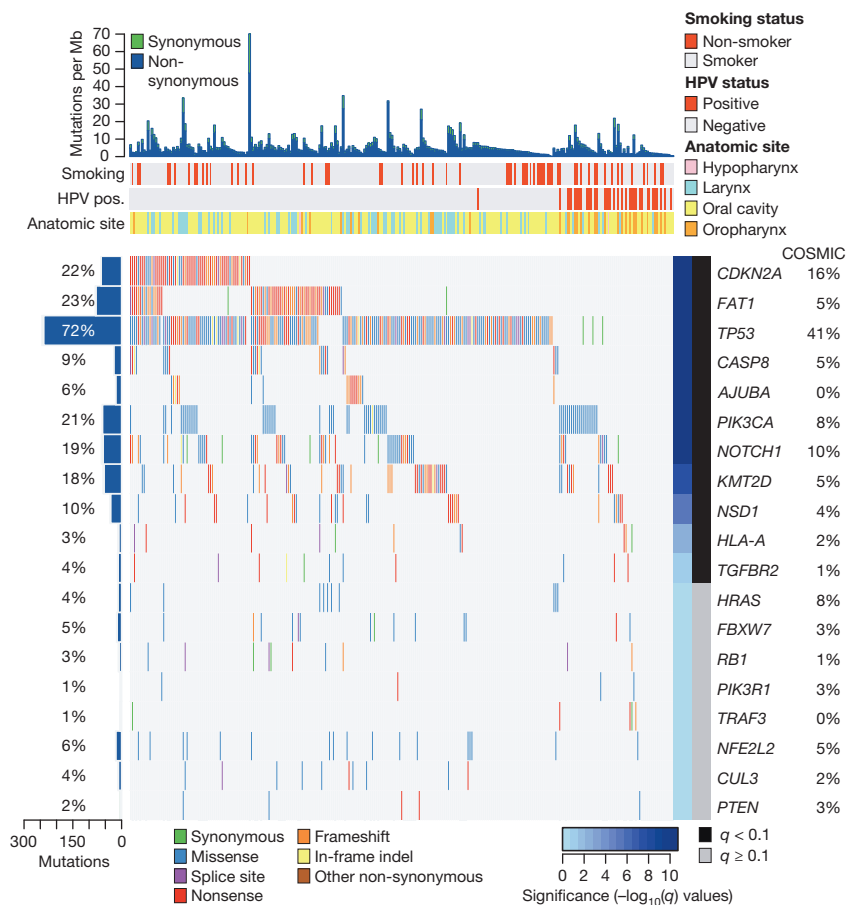


Figure 2 | Significantly mutated genes in HNSCC. Genes (rows) with significantly mutated genes (identified using the MutSigCV algorithm; $q < 0.1$) ordered by q value; additional genes with trends towards significance are also shown. Samples (columns, $n = 279$) are arranged to emphasize mutual exclusivity among mutations. Left, mutation percentage in TCGA. Right, mutation percentage in COSMIC ('upper aerodigestive tract' tissue). Top, overall number of mutations per megabase. Colour coding indicates mutation type.

centrosomal protein that regulates cell division, vertebrate ciliogenesis and left-right axis determination²³. Additionally, *AJUBA* is subject to EGFR-RAS-MAPK-dependent phosphorylation and implicated in Hippo growth and regeneration pathways conserved from *Drosophila* to mammals^{24,25}, in ataxia-telangiectasia mutated (ATM) and ATM and Rad-3-related (ATR)-mediated DNA damage response²⁶, and tumour invasion and migration²⁷.

A frequently mutated novel gene, the nuclear receptor binding SET domain protein 1 (*NSD1*), was identified in 33 HNSCCs. Alterations included inactivating mutations ($n = 29$) and focal homozygous deletions ($n = 4$). *NSD1* is a histone 3 Lys 36 (H3K36) methyltransferase, similar to *SETD2*, which is frequently mutated in the clear cell variant of renal cell carcinoma, and associated with DNA hypomethylation²⁸. Germline carriers of inactivating mutations in *NSD1* are associated with craniofacial abnormalities (Sotos syndrome), and malignancies including squamous carcinoma, implicating *NSD1* as a tumour suppressor gene²⁹. Interestingly, *NSD1* functions as an oncogene when fused to nucleoporin-98 (*NUP98*) t(5;11)(q35;p15.5) in haematological cancers with increased H3K36 trimethylation levels at *HOXA* genes and accompanying transcriptional activation³⁰. Translocations involving other dedicated H3K36 methyltransferase genes including *WHSC1* (also known as *MMSET* and *NSD2*) are reported in 20% of multiple myelomas. By contrast, *NSD1* loss has been associated with sporadic non-melanoma skin cancers³¹. Significant inactivating mutations were found in genes linked to squamous differentiation including in *NOTCH1* (19%), and other non-significant family members (*NOTCH2* 9%, and *NOTCH3* 5%, $q > 0.1$, non-significant), and the *TP63* target gene *ZNF750* (4%, $q > 0.1$, non-significant), which falls in a significantly deleted peak at 17q25.3. The analysis identified additional mutations including *TRAF3*, *RB1* and *NFE2L2*, among others with q values < 1 (non-significant). The frequently mutated apoptosis gene *CASP8* displayed clustered missense and other inactivating mutations in the first death effector,

intron and caspase peptidase domains. Statistically significant mutations in *KMT2D* (also known as *MLL2*) and *HLA-A* could contribute to defective immunosurveillance. Of known oncogenes, only *PIK3CA* achieved statistical significance ($q < 0.01$). Approximately one-quarter of the mutated *PIK3CA* cases displayed concurrent amplification, with an additional 20% of tumours containing focal amplification without evidence of mutation. Seventy-three per cent of *PIK3CA* mutations localized to Glu542Lys, Glu545Lys and His1047Arg/Leu hotspots that promote activation, with the remaining mutations of uncertain function. Recurrent activating mutations of *HRAS* in the GTPase domain in residues 11–13 approached statistical significance ($q = 0.2$).

We extended our unsupervised genome-wide analysis of significantly mutated genes as well as genes reported in COSMIC to a subgroup analysis by anatomic sites, tumour versus normal status, HPV status and four previously validated gene expression subtypes^{32,33} (Supplementary Information section 5, Supplementary Figs 5.1–5.4 and Supplementary Data 4.1 and 5.1–5.4). Additional mutations included *TRAF3*, *RB1* and *NFE2L2*, among others with q values < 1 , and we observed statistical evidence for mutations of *HRAS* ($q = 0$ in COSMIC subset) and other genes. Sporadic inactivating mutations and deletions of *TGFB2* were identified primarily in oral cavity tumours, consistent with its role in promoting squamous tumorigenesis in mouse models³⁴. Investigating COSMIC database mutations focused attention on the significant deletion peak at 4q31.3 containing the gene *FBXW7*, a ubiquitin ligase targeting cyclin E and *NOTCH* genes, in which we identified mutations that included recurrent Arg505Gly/Leu substitutions ($n = 14$). Genes with at least one identical mutation previously reported in COSMIC include *SCN9A*, *CHEK2*, *PTCH1* and *PIK3R1*. We further focused on somatic alterations and protein expression that represent plausible therapeutic targets (Fig. 3, Supplementary Information sections 6 and 7, Supplementary Figs 6.1 and 6.2 and Supplementary Data 6.1 and 6.2).

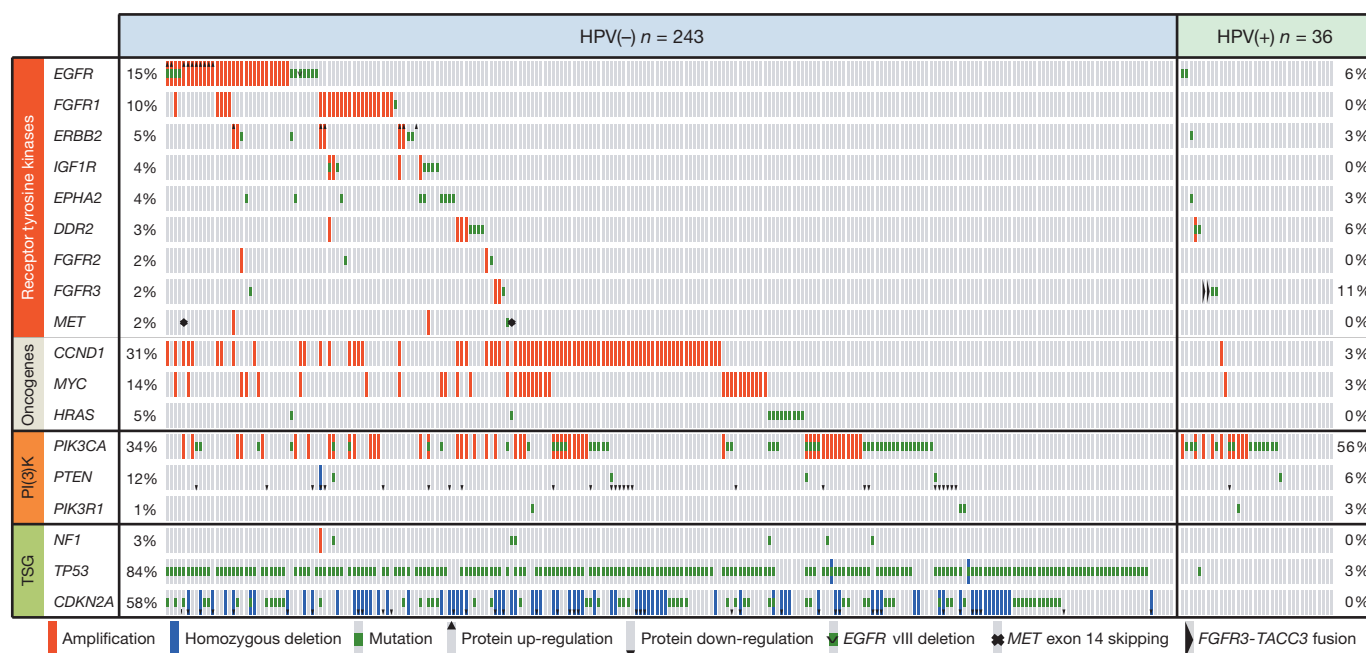


Figure 3 | Candidate therapeutic targets and driver oncogenic events. Alteration events for key genes are displayed by sample ($n = 279$). TSG, tumour suppressor gene.

Integrated genome analysis and pathways

Correlative genetic alteration analysis identified numerous pairwise significant findings (Supplementary Information section 7 and Supplementary Fig. 7.1). In particular, co-amplification of 11q13 containing *CCND1*, *FADD* and *CTTN* and a narrow segment of 11q22 containing the genes with equal evidence for *YAP1* and *BIRC2* was further characterized (Supplementary Fig. 7.2). Chromosome 11q22 was focally but rarely amplified in the absence of co-amplification of 11q13. This novel finding suggests that the selection pressure for this co-amplification stems from the interaction of *BIRC2* with *FADD* and the caspase cascade that inhibits cell death. Notably, the vast majority of tumours with the 11q13 amplification had large deletions in the telomeric region of 11q22, including other genes known to be important in cell death in cancer such as *ATM* and *CASP1*, 4, 5 and 12. Amplification of 11q13 was anti-correlated with *CASP8* mutations, suggesting an alternative function of *CASP8* and *FADD* in cell death/NF- κ B activation³⁵.

We investigated whether clinical factors, single gene alterations and statistically significant pairwise gene correlations (Supplementary Fig. 7.1 and Supplementary Data 7.1) might segregate previously defined molecular subtypes and/or anatomic sub-sites (Supplementary Data 5.1–5.4). We confirmed reported gene expression subtypes (atypical (24%), mesenchymal (27%), basal (31%) and classical (18%)), and assessed the subtypes for enrichment of somatic alterations^{32,33} (Supplementary Data 7.1). Notably, *TP53* mutation, *CDKN2A* loss of function, chromosome 3q amplification, alteration of oxidative stress genes (*KEAP1*, *NFE2L2* or *CUL3*), heavy smoking history (Supplementary Table 1.1) and larynx sub-site co-occurred in most classical subtype tumours (Fig. 4a and Supplementary Information section 7.2), similar to LUSC⁵ (Supplementary Figs 5.1 and 5.2). Collectively, these findings suggest that the *NFE2L2* oxidative stress pathway is a tobacco-related signature across anatomic tumour sites. By contrast, the basal subtype demonstrated inactivation of *NOTCH1* with intact oxidative stress

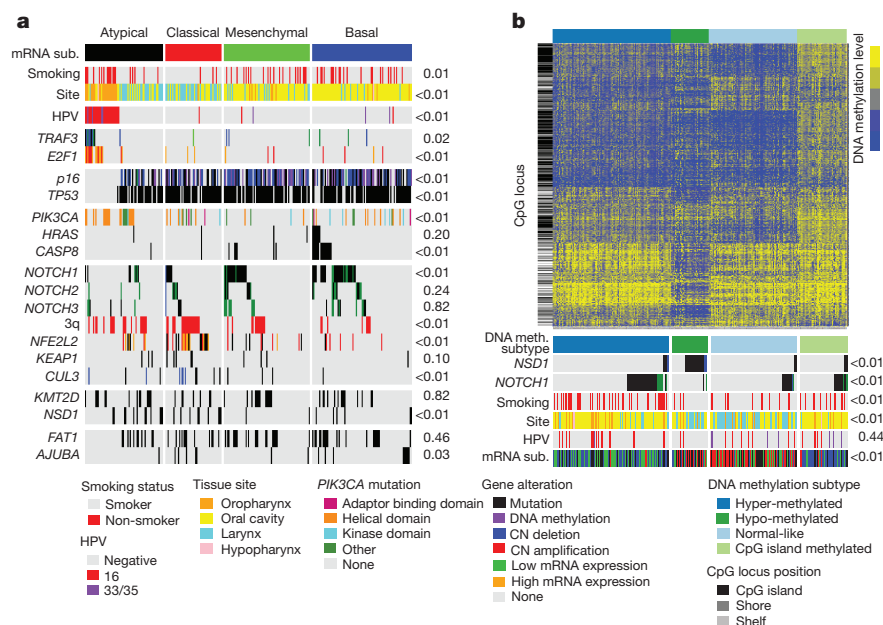


Figure 4 | Integrated analysis of genomic alterations. **a**, **b**, Samples ($n = 279$) are displayed in columns and grouped by gene expression (**a**) or methylation (**b**) subtype (sub.). Unadjusted two-sided Fisher's exact test *P* values assess the association of each genomic alteration. Methylation probe location of CpG islands, shores and shelves are shown on the left of **b**. Annotation shows HPV status and subtype (16, 33 and 35). CN, copy number.

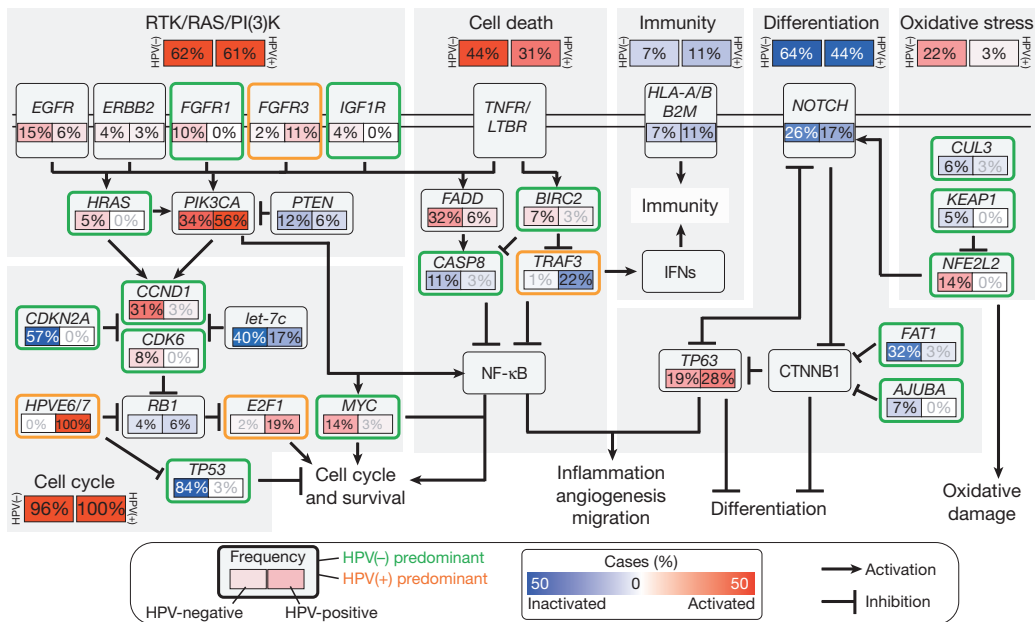


Figure 5 | Deregulation of signalling pathways and transcription factors. Key affected pathways, components and inferred functions, are summarized in the main text and Supplementary Information section 7 for $n = 279$ samples. The frequency (%) of genetic alterations for HPV(-) and HPV(+) tumours are shown separately within sub-panels and highlighted. Also see Supplementary Fig. 7.15. Pathway alterations include homozygous deletions, focal amplifications and somatic mutations. Activated and inactivated pathways/genes, and activating or inhibitory symbols are based on predicted effects of genome alterations and/or pathway functions.

signalling and fewer alterations of chromosome 3q. Analysis of the 3q locus highlighted a marked relative decrease of *SOX2* expression in basal tumours relative to all other HNSCC and tumour adjacent normal samples (Supplementary Fig. 5.4), supporting the interaction of transcription factors *SOX2*, *TP63*, *NFE2L2* and *NOTCH1* as driving differences between expression subtypes. Additionally, the basal subtype included most tumours with the *HRAS*–*CASP8* co-mutation and most co-amplified 11q13/q22 tumours. These findings along with *HRAS* mutations implicate disrupted cell death as a major alteration in this subtype³⁶ (Supplementary Fig. 7.2). The atypical subtype was characterized by a lack of chromosome 7 amplifications (Supplementary Fig. 5.3), enrichment of HPV(+) tumours with activating mutations in exon 9 that contains the *PIK3CA* helical domain. By contrast, the mesenchymal subtype showed high levels of alteration in innate immunity genes, in particular high expression of natural killer cell marker *CD56* and a low frequency of HLA class I mutations (Supplementary Fig. 7.3). Among the significantly mutated genes, *TP53* ($P < 0.001$), *CASP8* ($P = 0.01$), *NSD1* ($P = 0.01$) and *CDKN2A* (0.06) were the most differentially mutated across anatomic sites (Supplementary Data 4.1). Most *CASP8* mutations (22 out of 24, 92%) were in oral cavity tumours, whereas *TP53*, *NSD1* and *CDKN2A* demonstrated decreased mutation rates in oropharyngeal tumours relative to other sites.

Unsupervised analysis of gene expression by HPV status and of reverse-phase protein arrays (Supplementary Information section 6), DNA methylation (Supplementary Information section 8), and miRNA platforms (Supplementary Information section 9, Supplementary Table 9.1, Supplementary Figs 7.4–7.9 and Supplementary Data 7.1) showed high correlation across platforms ($P < 0.01$; Fig. 4, Supplementary Information section 7.9 and Supplementary Data 7.2) and coordinated alterations of genes including the epithelial–mesenchymal transition signature³⁷ (Supplementary Figs 7.4, 7.7 and 7.8). However, within the broader cross-platform agreement, individual unsupervised clustering of miRNA, reverse-phase protein arrays and DNA methylation data provided insight into the association of molecular subtypes with single gene alterations. The most notable example was the detection of hypomethylation and loss-of-function mutations of *NSD1*, and wild-type *NOTCH1* in atypical and classical gene expression subtypes (Fig. 4b and Supplementary Table 7.1).

Supervised analyses detected genomic features (miRNA, gene expression and DNA methylation) associated with anatomic site (Supplementary Figs 5.5–5.8 and Supplementary Data 5.1–5.4). A supervised integrated analysis identified target genes that are inversely regulated by

miRNAs in HNSCC (Supplementary Information section 7.4). Among these miRNA–messenger RNA networks, let-7c-5p and miR-100-5p exhibited a correlation between low copy number and expression. Let-7c-5p and miR-100-5p were decreased in tumours compared to normal (Supplementary Fig. 7.10). For these miRNAs, deletion was highly associated with increased expression of target genes, including the cell cycle regulator *CDK6*, transcription factor *E2F1* (ref. 38), mitosis regulator *PLK1* (ref. 39), and transcription factor *HMGA2* (ref. 40; Supplementary Figs 7.10, Supplementary Tables 7.2 and 7.3).

Integrative bioinformatics analysis identified a limited number of pathways targeted by frequent genome alterations (Fig. 5, Supplementary Information section 7, Supplementary Figs 7.11–7.15 and Supplementary Data 7.3). Among receptor tyrosine kinases, *EGFR/ERBB2* or *FGFR1/3* alterations are the most frequent. Among downstream targets of the receptor tyrosine kinase (RTK)/RAS/phosphatidylinositol-3-OH kinase (PI(3)K) pathway, *PIK3CA* dominates with occasional *HRAS* and *PTEN* alterations. Further downstream, nearly every tumour has alteration of genes governing the cell cycle. The tumour suppressors *TP53* and *CDKN2A*, oncogenes *CCND1* and *MYC*, and the newly identified miRNA let-7c, are most often altered in HPV(-) tumours, whereas viral genes *E6*, *E7* and *E2F1* predominate in HPV(+) cases. In addition, we report frequent alterations in genes involved in cell death, NF- κ B-mediated survival, or immunity pathways^{15,35}. Co-amplification of *FADD* \pm *BIRC2*, or *CASP8* \pm *HRAS* mutations define exclusive HPV(-) subsets, whereas *TRAF3* loss characterizes an HPV(+) subset. These alterations along with *PIK3CA* and *TP63* converge on NF- κ B transcription factors that promote cell survival, migration, inflammation and angiogenesis^{41,42}. Furthermore, *TRAF3* and/or *HLA* loss are implicated in deregulation of innate antiviral and adaptive anti-tumour immunity^{43,44}. Further alterations of *NOTCH*, *TP63* and other genes in HPV(-) tumours (*FAT1* and *AJUBA*) recently linked functionally to β -catenin (*CTNNB1*) are also detected^{21,22,45}. Finally, we highlight a previously underappreciated role for a key transcription factor regulator of oxidative stress, *NFE2L2*, and its protein complex partners *CUL3* and *KEAP1* in HPV(-) HNSCCs.

Conclusion

The TCGA study represents the most comprehensive integrative genomic analysis of HNSCC. Loss of *TRAF3*, activating mutations of *PIK3CA*, and amplification of *E2F1* in HPV(+) oropharyngeal cancers point to aberrant activation of NF- κ B, other oncogenic pathways, and cell cycle, as critical in the pathogenesis and development of new

targeted therapies for these tumours. In HPV(–) HNSCCs, mutually exclusive subsets containing amplicons on 11q with *CCND1*, *FADD*, *BIRC2* and *YAP1*, or concurrent mutations of *CASP8* with *HRAS*, also target cell cycle, death, NF- κ B and other oncogenic pathways. Recent studies predict that the inactivation of *AJUBA*, as well as *FAT1* and *NOTCH1*, may converge to uncheck Wnt/ β -catenin signalling, implicated in deregulation of cell polarity and differentiation. The 3q amplicon found in both HPV(+) and (–) HNSCCs includes transcription factors *TP63*, *SOX2* and signal molecule *PIK3CA*, which are also implicated in homeostasis of epithelial stem cells and differentiation. Among these, the biological function and agents targeting BIRCs, PI(3)K, Wnt/ β -catenin and *NOTCH* are under investigation. Collectively, these findings provide new insights into HNSCC and suggest that shared and unique alterations might be leveraged to accelerate progress in prevention and therapy across tumour types.

Received 19 October 2013; accepted 1 December 2014.

1. Ferlay, J. *et al.* Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* **127**, 2893–2917 (2010).
2. Ang, K. K. *et al.* Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* **363**, 24–35 (2010).
3. Stransky, N. *et al.* The mutational landscape of head and neck squamous cell carcinoma. *Science* **333**, 1157–1160 (2011).
4. Agrawal, N. *et al.* Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in *NOTCH1*. *Science* **333**, 1154–1157 (2011).
5. The Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525 (2012).
6. Oganessian, G. *et al.* Critical role of TRAF3 in the Toll-like receptor-dependent and -independent antiviral response. *Nature* **439**, 208–211 (2006).
7. Karim, R. *et al.* Human papillomavirus (HPV) upregulates the cellular ubiquitinase UCHL1 to suppress the keratinocyte's innate immune response. *PLoS Pathog.* **9**, e1003384 (2013).
8. Eliopoulos, A. G. *et al.* CD40-induced growth inhibition in epithelial cells is mimicked by Epstein–Barr Virus-encoded LMP1: involvement of TRAF3 as a common mediator. *Oncogene* **13**, 2243–2254 (1996).
9. Imbeault, M. *et al.* Acquisition of host-derived CD40L by HIV-1 *in vivo* and its functional consequences in the B-cell compartment. *J. Virol.* **85**, 2189–2200 (2011).
10. Ni, C. Z. *et al.* Molecular basis for CD40 signaling mediated by TRAF3. *Proc. Natl Acad. Sci. USA* **97**, 10395–10399 (2000).
11. Chung, G. T. *et al.* Constitutive activation of distinct NF- κ B signals in EBV-associated nasopharyngeal carcinoma. *J. Pathol.* **231**, 311–322 (2013).
12. Annunziata, C. M. *et al.* Frequent engagement of the classical and alternative NF- κ B pathways by diverse genetic abnormalities in multiple myeloma. *Cancer Cell* **12**, 115–130 (2007).
13. Ciriello, G. *et al.* Emerging landscape of oncogenic signatures across human cancers. *Nature Genet.* **45**, 1127–1133 (2013).
14. Smeets, S. J. *et al.* Genetic classification of oral and oropharyngeal carcinomas identifies subgroups with a different prognosis. *Cellular Oncol.* **31**, 291–300 (2009).
15. Mayo, M. W. *et al.* Requirement of NF- κ B activation to suppress p53-independent apoptosis induced by oncogenic Ras. *Science* **278**, 1812–1815 (1997).
16. Sok, J. C. *et al.* Mutant epidermal growth factor receptor (EGFRvIII) contributes to head and neck cancer growth and resistance to EGFR targeting. *Clin. Cancer Res.* **12**, 5064–5073 (2006).
17. Kong-Beltran, M. *et al.* Somatic mutations lead to an oncogenic deletion of met in lung cancer. *Cancer Res.* **66**, 283–289 (2006).
18. Popescu, N. C., DiPaolo, J. A. & Amsbaugh, S. C. Integration sites of human papillomavirus 18 DNA sequences on HeLa cell chromosomes. *Cytogenet. Cell Genet.* **44**, 58–62 (1987).
19. Forbes, S. A. *et al.* The Catalogue of Somatic Mutations in Cancer (COSMIC). *Curr. Protoc. Hum. Genet.* Chapter 10, Unit 10.11 (2008).
20. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
21. Morris, L. G. *et al.* Recurrent somatic mutation of *FAT1* in multiple human cancers leads to aberrant NF- κ B activation. *Nature Genet.* **45**, 253–261 (2013).
22. Haraguchi, K. *et al.* Ajuba negatively regulates the Wnt signaling pathway by promoting GSK-3 β -mediated phosphorylation of β -catenin. *Oncogene* **27**, 274–284 (2008).
23. Nagai, Y. *et al.* The LIM protein Ajuba is required for ciliogenesis and left-right axis determination in medaka. *Biochem. Biophys. Res. Commun.* **396**, 887–893 (2010).
24. Sun, G. & Irvine, K. D. Ajuba family proteins link JNK to Hippo signaling. *Sci. Signal.* **6**, ra81 (2013).
25. Reddy, B. V. & Irvine, K. D. Regulation of Hippo signaling by EGFR-MAPK signaling through Ajuba family proteins. *Dev. Cell* **24**, 459–471 (2013).
26. Kalan, S., Matveyenko, A. & Loayza, D. LIM protein Ajuba participates in the repression of the ATR-mediated DNA damage response. *Front. Genet.* **4**, 95 (2013).
27. Nola, S. *et al.* Ajuba is required for Rac activation and maintenance of E-cadherin adhesion. *J. Cell Biol.* **195**, 855–871 (2011).
28. The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **499**, 43–49 (2013).
29. Fickie, M. R. *et al.* Adults with Sotos syndrome: review of 21 adults with molecularly confirmed *NSD1* alterations, including a detailed case report of the oldest person. *Am. J. Med. Genet. A* **155**, 2105–2111 (2011).
30. Wang, G. G., Cai, L., Pasillas, M. P. & Kamps, M. P. NUP98–NSD1 links H3K36 methylation to *Hox-A* gene activation and leukaemogenesis. *Nature Cell Biol.* **9**, 804–812 (2007).
31. Quintana, R. M. *et al.* A transposon-based analysis of gene mutations related to skin cancer development. *J. Invest. Dermatol.* **133**, 239–248 (2013).
32. Chung, C. H. *et al.* Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. *Cancer Cell* **5**, 489–500 (2004).
33. Walter, V. *et al.* Molecular subtypes in head and neck cancer exhibit distinct patterns of chromosomal gain and loss of canonical cancer genes. *PLoS ONE* **8**, e56823 (2013).
34. Lu, S. L. *et al.* Loss of transforming growth factor- β type II receptor promotes metastatic head-and-neck squamous cell carcinoma. *Genes Dev.* **20**, 1331–1342 (2006).
35. Oberst, A. & Green, D. R. It cuts both ways: reconciling the dual roles of caspase 8 in cell death and survival. *Nature Rev. Mol. Cell Biol.* **12**, 757–763 (2011).
36. Park, S. J. *et al.* Opposite role of Ras in tumor necrosis factor- α -induced cell cycle regulation: competition for Raf kinase. *Biochem. Biophys. Res. Commun.* **287**, 1140–1147 (2001).
37. Byers, L. A. *et al.* An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. *Clin. Cancer Res.* **19**, 279–290 (2013).
38. Wong, J. V., Dong, P., Nevins, J. R., Mathey-Prevot, B. & You, L. Network calisthenics: control of E2F dynamics in cell cycle entry. *Cell Cycle* **10**, 3086–3094 (2011).
39. Sanhaji, M. *et al.* Polo-like kinase 1 inhibitors, mitotic stress and the tumor suppressor p53. *Cell Cycle* **12**, 1340–1351 (2013).
40. Morishita, A. *et al.* HMGA2 is a driver of tumor metastasis. *Cancer Res.* **73**, 4289–4299 (2013).
41. Bancroft, C. C. *et al.* Effects of pharmacologic antagonists of epidermal growth factor receptor, PI3K and MEK signal kinases on NF- κ B and AP-1 activation and IL-8 and VEGF expression in human head and neck squamous cell carcinoma lines. *Int. J. Cancer* **99**, 538–548 (2002).
42. Yang, X. *et al.* Δ Np63 versatily regulates a broad NF- κ B gene program and promotes squamous epithelial proliferation, migration, and inflammation. *Cancer Res.* **71**, 3688–3700 (2011).
43. Keating, P. J. *et al.* Frequency of down-regulation of individual HLA-A and -B alleles in cervical carcinomas in relation to TAP-1 expression. *Br. J. Cancer* **72**, 405–411 (1995).
44. Esteban, F. *et al.* Lack of MHC class I antigens and tumour aggressiveness of the squamous cell carcinoma of the larynx. *Br. J. Cancer* **62**, 1047–1051 (1990).
45. Dotto, G. P. Crosstalk of Notch with p53 and p63 in cancer growth control. *Nature Rev. Cancer* **9**, 587–595 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements This study was supported by National Institutes of Health (NIH) grants: P50CA097190, P50CA16672, U54 HG003273, U54 HG003067, U54 HG003079, U24 CA143799, U24 CA143835, U24 CA143840, U24 CA143843, U24 CA143845, U24 CA143848, U24 CA143858, U24 CA143866, U24 CA143867, U24 CA143882, U24 CA143883, U24 CA144025 and R01 CA095419. Supported by The Bobby F. Garrett Fund for Head and Neck Cancer Research and NIDCD Intramural Projects ZIA-DC-000016, 73 and 74.

Author Contributions The TCGA research network contributed collectively to this study. Biospecimens were provided by the tissue source sites and processed by the biospecimen core resource. Data generation and analyses were performed by the genome sequencing centres, cancer genome characterization centres and genome data analysis centres. All data were released through the data coordinating centre. Project activities were coordinated by the National Cancer Institute and National Human Genome Research Institute project teams. We also acknowledge the following TCGA investigators who made substantial contributions to the project: C.V.W. (manuscript coordinator); V.W. (data coordinator); P.S.H. and D.N.H. (analysis coordinators); Y.W. (data coordinating centre representative); R.A., J.B., T.E.C., J.J.L., M.Parfenov, A.D.T., V.W. and M.D.W. (clinical and HPV analysis); A.D.C. (DNA copy number analysis); J.Cho, A.H., M.Parfenov and M.D.W. (mutation calling); L.C. and L.Danilova (DNA methylation analysis); R.Bowlby, J.P.B., Z.C., H.Cheng, A.H., F.F.L., A.G.R., A.D.S. and C.V.W. (miRNA sequence analysis); L.A.B., L.Diao, P.K., W.L., S.N. and J.W. (reverse-phase protein arrays analysis); L.A.B., L.Diao, S.R.J., P.K.K., V.W., M.D.W. and N.Z. (mRNA sequence analysis); S.Benz, Z.C., R.L.F., S.N., C.R.P., N.Schultz, T.Y.S., C.V.W. and V.W. (pathway analysis); A.E.N., J.R.G. and D.N.H. (project chairs).

Author Information The primary and processed data used to generate the analyses can be downloaded by registered users from The Cancer Genome Atlas (<https://tcga-data.nci.nih.gov/tcga/tcgaDownload.jsp>, <https://cghub.ucsc.edu/> and https://tcga-data.nci.nih.gov/docs/publications/hnsc_2014/). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.N.H. (hayes@med.unc.edu), J.R.G. (grandisjr@upmc.edu) or A.E.N. (anaggar@mdanderson.org).



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported licence. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons licence, users will need to obtain permission from the licence holder

to reproduce the material. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-sa/3.0>

The Cancer Genome Atlas Research Network (Participants are arranged by area of contribution and then by institution.)

Genome sequencing centre: Broad Institute Michael S. Lawrence¹, Carrie Sougnez¹, Lee Lichtenstein¹, Kristian Cibulskis¹, Eric Lander¹, Stacey B. Gabriel¹, Gad Getz^{1,2}

Genome characterization and data analysis centres: BC Cancer Agency Adrian Ally³, Miruna Balasundaram³, Inanc Birol³, Reanne Bowlby³, Denise Brooks³, Yaron S. N. Butterfield³, Rebecca Carlsen³, Dean Cheng³, Andy Chu³, Noreen Dhalla³, Ranabir Guin³, Robert A. Holt³, Steven J. M. Jones³, Darlene Lee³, Haiyan I. Li³, Marco A. Marra³, Michael Mayo³, Richard A. Moore³, Andrew J. Mungall³, A. Gordon Robertson³, Jacqueline E. Schein³, Payal Sipahimalani³, Angela Tam³, Nina Thiessen³, Tina Wong³; **Harvard Medical School/Brigham & Women's Hospital/MD Anderson Cancer Center** Alexei Protopopov⁴, Netty Santoso⁵, Semin Lee⁵, Michael Parfenov⁵, Jianhua Zhang⁴, Harshad S. Mahadeshwar⁴, Jiabin Tang⁴, Xiaojia Ren⁵, Sahil Seth⁴, Psalm Haseley⁵, Dong Zeng⁴, Lixing Yang⁵, Andrew W. Xu⁵, Xingzhi Song⁴, Angeliki Pantazi⁵, Christopher A. Bristow⁴, Angela Hadjipanayis^{5,6}, Jonathan Seidman⁷, Lynda Chin^{1,4}, Peter J. Park^{5,6,7}, Raju Kucheralapati^{5,6}; **The University of Texas MD Anderson Cancer Center** Rehan Akbani⁴, Tod Casasent⁴, Wenbin Liu⁴, Yiling Lu⁴, Gordon Mills⁴, Thomas Motter⁴, John Weinstein⁴, Lixia Dia⁴, Jing Wang⁴, You Hong Fan⁴; **University of Kentucky** Jinze Liu⁸, Kai Wang⁸; **University of North Carolina at Chapel Hill** J. Todd Auman⁹, Saianand Balu⁹, Thomas Bodenheimer⁹, Elizabeth Buda⁹, D. Neil Hayes⁹, Katherine A. Hoadley⁹, Alan P. Hoyle⁹, Stuart R. Jefferys⁹, Corbin D. Jones⁹, Patrick K. Kimes⁹, Yufeng Liu⁹, J. S. Marron⁹, Shaowu Meng⁹, Piotr A. Mieczkowski⁹, Lisle E. Mose⁹, Joel S. Parker⁹, Charles M. Perou⁹, Jan F. Prins⁹, Jeffrey Roach⁹, Yan Shi⁹, Janae V. Simons⁹, Darshan Singh⁹, Matthew G. Soloway⁹, Donghui Tan⁹, Umadevi Veluvolu⁹, Vonn Walter⁹, Scot Waring⁹, Matthew D. Wilkerson⁹, Junyuan Wu⁹, Ni Zhao⁹; **Broad Institute** Andrew D. Cherniack¹, Peter S. Hammerman^{1,10}, Aaron D. Tward^{1,5}, Chandra Sekhar Pedamallu^{1,10}, Gordon Saksena¹, Joonil Jung¹, Akinyemi I. Ojesina^{1,10}, Scott L. Carter¹, Travis I. Zack^{1,10,11}, Steven E. Schumacher^{1,10}, Rameen Beroukhi^{1,10}, Samuel S. Freeman¹, Matthew Meyerson^{1,5,10}, Juok Cho¹, Lynda Chin^{1,4}, Gad Getz^{1,2}, Michael S. Noble¹, Daniel DiCara¹, Hailei Zhang¹, David I. Heiman¹, Nils Gehlenborg^{1,5}, Doug Voelt¹, Pei Lin¹, Scott Frazer¹, Petar Stojanov^{1,10}, Yingchun Liu¹, Lihua Zou¹, Jaegil Kim¹, Carrie Sougnez¹, Stacey B. Gabriel¹, Michael S. Lawrence¹; **Baylor College of Medicine** Donna Muzny^{1,2}, HarshaVardhan Doddapaneni^{1,2}, Christie Kovar^{1,2}, Jeff Reid^{1,2}, Donna Morton^{1,2}, Yi Han^{1,2}, Walker Hale^{1,2}, Hsu Chao^{1,2}, Kyle Chang^{1,2}, Jennifer A. Drummond^{1,2}, Richard A. Gibbs^{1,2}, Nipun Kakkar^{1,2}, David Wheeler^{1,2}, Liu Xi^{1,2}; **Memorial Sloan-Kettering Cancer Center** Giovanni Ciriello^{1,3}, Marc Ladanyi^{1,3}, William Lee^{1,3}, Ricardo Ramirez^{1,3}, Chris Sander^{1,3}, Ronglai Shen^{1,3}, Rileen Sinha^{1,3}, Nils Wernhold^{1,3}, Barry S. Taylor^{1,3}, B. Arman Aksoy^{1,3}, Gideon Dresdner^{1,3}, Jianjiong Gao^{1,3}, Benjamin Gross^{1,3}, Anders Jacobsen^{1,3}, Boris Reva^{1,3}, Nikolaus Schultz^{1,3}, S. Onur Sumner^{1,3}, Yichao Sun^{1,3}, Timothy A. Chan^{1,3}, Luc G. Morris^{1,3}; **University of California Santa Cruz/Buck Institute** Joshua Stuart^{1,4}, Stephen Benz^{1,4}, Sam Ng^{1,4}, Christopher Benz^{1,4}, Christina Yau^{1,4}; **Johns Hopkins University/Sidney Kimmel Comprehensive Cancer Center** Stephen B. Baylin^{1,5}, Leslie Cope^{1,5}, Ludmila Danilova^{1,5}, James G. Herman^{1,5}; **University of Southern California** Moiz Bootwalla^{1,6}, Dennis T. Maglinte^{1,6}, Peter W. Laird^{1,6}, Timothy Triche Jr^{1,6}, Daniel J. Weisenberger^{1,6}, David J. Van Den Berg^{1,6}

Disease working group: Nishant Agrawal^{1,7}, Justin Bishop^{1,7}, Paul C. Boutros^{1,8}, Jeff P. Bruce^{1,9}, Lauren Averett Byers⁴, Joseph Califano^{1,7}, Thomas E. Carey²⁰, Zhong Chen²¹, Hui Cheng²¹, Simion I. Chiosea²², Ezra Cohen²³, Brenda Diergaarde²², Ann Marie Egloff²², Adel K. El-Naggar⁴, Robert L. Ferris²², Mitchell J. Frederick⁴, Jennifer R. Grandis²², Yan Guo²⁴, Robert I. Haddad¹⁰, Peter S. Hammerman^{1,10}, Thomas Harris²⁵, D. Neil Hayes⁹, Angela B. Y. Hui¹⁹, J. Jack Lee⁴, Scott M. Lippman²⁶, Fei-Fei Liu²⁷, Jonathan B. McHugh²⁰, Jeff Myers⁴, Patrick Kwok Shing Ng⁴, Bayardo Perez-Ordenez²⁸, Curtis R. Pickering⁴, Michael Prystowsky²⁵, Marjorie Romkes²², Anthony D. Saleh²¹, Maureen A. Sartor²⁰, Raja Seethala²², Tangy U. Seiwer²³, Han Si²¹, Aaron D. Tward^{1,5}, Carter Van Waes²¹, Daryl M. Waggott¹⁸, Maciej Wiznerowicz²⁹, Wendell G. Yarbrough³⁰, Jiexin Zhang⁴, Zhixiang Zuo²³

Biospecimen core resource: International Genomics Consortium Ken Burnett³¹, Daniel Crain³¹, Johanna Gardner³¹, Kevin Lau³¹, David Mallery³¹, Scott Morris³¹, Joseph Paulauskis³¹, Robert Penny³¹, Candace Shelton³¹, Troy Shelton³¹, Mark Sherman³¹, Peggy Yena³¹; **Nationwide Children's Hospital** Aaron D. Black³², Jay Bowen³², Jessica Frick³², Julie M. Gastier-Foster³², Hollie A. Harper³², Kristen Leraas³², Tara M. Lichtenberg³², Nilsa C. Ramirez³², Lisa Wise³², Erik Zmuda³²

Data coordinating centre: Julien Baboud³³, Mark A. Jensen³³, Ari B. Kahn³³, Todd D. Pihl³³, David A. Pot³³, Deepak Srinivasan³³, Jessica S. Walton³³, Yunhu Wan³³

Project office: Robert A. Burton³⁴, Tanja Davidsson³⁵, John A. Demchok³⁵, Greg Eley³⁶, Martin L. Ferguson³⁵, Kenna R. Mills Shaw³⁵, Bradley A. Ozenberger³⁷, Margi Sheth³⁵, Heidi J. Sofia³⁷, Roy Tarnuzzer³⁵, Zhining Wang³⁵, Liming Yang³⁵, Jean Claude Zenklusen³⁵

Tissue source sites: Analytical Biological Services Charles Saller³⁸, Katherine Tarvin³⁸; **Fred Hutchinson Cancer Research Center** Chu Chen³⁹; **Georgia Regents University** Roni Bollag⁴⁰, Paul Weinberger⁴⁰; **Greater Poland Cancer Centre** Wojciech Golusiński²⁹, Pawel Golusiński²⁹, Matthew Ibbs²⁹, Konstanty Korski²⁹, Andrzej Mackiewicz²⁹, Wiktoria Suchorska²⁹, Bartosz Szybiak²⁹, Maciej Wiznerowicz²⁹; **International Genomics Consortium** Ken Burnett³¹, Erin Curley³¹, Johanna Gardner³¹, David Mallery³¹, Robert Penny³¹, Troy Shelton³¹, Peggy Yena³¹; **Indiana University Simon Cancer Center** Christina Beard⁴¹, Colleen Mitchell⁴¹, George Sandusky⁴¹; **Johns Hopkins University** Nishant Agrawal^{1,7}, Julie Ahn^{1,7}, Justin Bishop^{1,7}, Joseph Califano^{1,7}, Zubair Khan^{1,7}; **Princess Margaret Cancer Centre** Jeff P. Bruce¹⁹, Angela B. Y. Hui¹⁹, Jonathan Irish²⁷, Fei-Fei Liu²⁷, Bayardo Perez-Ordenez²⁸, John Waldron²⁷; **Ontario Institute for Cancer Research** Paul C. Boutros¹⁸, Daryl M. Waggott¹⁸; **The University of Texas MD Anderson Cancer Center** Jeff Myers⁴, William N. William Jr⁴; **University of California San Diego** Scott M. Lippman²⁶; **University of Miami** Sophie Egea⁴², Carmen Gomez-Fernandez⁴², Lynn Herbert⁴²; **University of Michigan** Carol R. Bradford²⁰, Thomas E. Carey²⁰, Douglas B. Chepeha²⁰, Andrea S. Haddad²⁰, Tamara R. Jones²⁰, Christine M. Komarck²⁰, Mayya Malakh²⁰, Jonathan B. McHugh²⁰, Jeffrey S. Moyer²⁰, Ariane Nguyen²⁰, Lisa A. Peterson²⁰, Mark E. Prince²⁰, Laura S. Rozek²⁰, Maureen A. Sartor²⁰, Evan G. Taylor²⁰, Heather M. Walline²⁰, Gregory T. Wolf²⁰; **University of North Carolina at Chapel Hill** Lori Boice⁹, Bishamjit S. Chera⁹, William K. Funkhouser⁹, Margaret L. Gulley⁹, Trevor G. Hackman⁹, D. Neil Hayes⁹, Michele C. Hayward⁹, Mei Huang⁹, W. Kimryn Rathmell⁹, Ashley H. Salazar⁹, William W. Shockley⁹, Carol G. Shores⁹, Leigh Thorne⁹, Mark C. Weissler⁹, Sylvia Wrenn⁹, Adam M. Zanation⁹; **University of Pittsburgh** Simion I. Chiosea²², Brenda Diergaarde²², Ann Marie Egloff²², Robert L. Ferris²², Marjorie Romkes²², Raja Seethala²²; **Vanderbilt University** Brandee T. Brown²⁴, Yan Guo²⁴, Michelle Pham²⁴; **Yale University** Wendell G. Yarbrough³⁰

¹The Eli and Edythe L. Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts 02142, USA. ²Massachusetts General Hospital, 55 Fruit Street, Boston, Massachusetts 02114, USA. ³Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia V5Z 4S6, Canada. ⁴The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. ⁵Harvard Medical School, Boston, Massachusetts 02115, USA. ⁶Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. ⁷Boston Children's Hospital, Boston, Massachusetts 02115, USA. ⁸University of Kentucky, Lexington, Kentucky 40506, USA. ⁹University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. ¹⁰Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, Massachusetts 02215, USA. ¹¹Harvard University, Boston, Massachusetts 02115, USA. ¹²Baylor College of Medicine, Houston, Texas 77030, USA. ¹³Memorial Sloan-Kettering Cancer Center, 1275 York Ave, New York, New York 10065, USA. ¹⁴University of California Santa Cruz, Santa Cruz, California 95064, USA. ¹⁵The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University, Baltimore, Maryland 21287, USA. ¹⁶University of Southern California, 1450 Biggy Street, NRT G511, Los Angeles, California 90033, USA. ¹⁷Johns Hopkins Medical Institutions, Baltimore, Maryland 21205, USA. ¹⁸Ontario Institute for Cancer Research 101 College Street, Suite 800, Toronto, Ontario M5G 0A3, Canada. ¹⁹Ontario Cancer Institute, Toronto, Ontario M5G 2M9, Canada. ²⁰University of Michigan Comprehensive Cancer Center, Ann Arbor, Michigan 48109, USA. ²¹National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Bethesda, Maryland 20892, USA. ²²University of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA. ²³The University of Chicago, Chicago, Illinois 60637, USA. ²⁴Vanderbilt University Medical Center 7209 Medical Center East, South Tower Nashville, Tennessee 37232, USA. ²⁵Albert Einstein College of Medicine, Bronx, New York 10461, USA. ²⁶University of California San Diego, San Diego, California 92093, USA. ²⁷University of Toronto, Toronto, Ontario M5G 2M9, Canada. ²⁸University Health Network, Toronto, Ontario M5G 2C4, Canada. ²⁹Greater Poland Cancer Centre, Garbary 15, 61-866, Poznań, Poland. ³⁰Yale School of Medicine, 800 Howard Avenue, 4th Floor Yale Physicians Building, New Haven, Connecticut 06520-8041, USA. ³¹International Genomics Consortium, 445 North 5th Street, Phoenix, Arizona 85004, USA. ³²The Research Institute at Nationwide Children's Hospital, Columbus, Ohio 43205, USA. ³³SRA International, Fairfax, Virginia 22033, USA. ³⁴Frederick National Laboratory for Cancer Research, 1050 Boyles Street, Frederick, Maryland 21702, USA. ³⁵National Cancer Institute, 31 Center Drive 3A20, Bethesda, Maryland 20892, USA. ³⁶Scimientis, LLC, Atlanta, Georgia 30666, USA. ³⁷National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. ³⁸Analytical Biological Services, Inc., 701 Cornell Drive, Wilmington, Delaware 19801, USA. ³⁹Fred Hutchinson Cancer Research Center, Mailstop M5-C800, 1100 Fairview Avenue North, Seattle, Washington 98109, USA. ⁴⁰Georgia Regents University, Augusta, Georgia 30912, USA. ⁴¹Indiana University Simon Cancer Center, Indianapolis, Indiana 46202, USA. ⁴²University of Miami, Sylvester Comprehensive Cancer Center, 1550 NW 10th Avenue, Miami, Florida 33136, USA.

Genome-scale transcriptional activation by an engineered CRISPR–Cas9 complex

Silvana Konermann^{1,2,3,4*}, Mark D. Brigham^{1,2,3,4*}, Alexandro E. Trevino^{1,2,3,4}, Julia Joung^{1,4}, Omar O. Abudayyeh^{1,2,3,4}, Clea Barcena^{1,2,3,4}, Patrick D. Hsu^{1,2,3,4}, Naomi Habib¹, Jonathan S. Gootenberg^{1,2,3,4,5}, Hiroshi Nishimasu^{6,7}, Osamu Nureki⁶ & Feng Zhang^{1,2,3,4}

Systematic interrogation of gene function requires the ability to perturb gene expression in a robust and generalizable manner. Here we describe structure-guided engineering of a CRISPR–Cas9 complex to mediate efficient transcriptional activation at endogenous genomic loci. We used these engineered Cas9 activation complexes to investigate single-guide RNA (sgRNA) targeting rules for effective transcriptional activation, to demonstrate multiplexed activation of ten genes simultaneously, and to upregulate long intergenic non-coding RNA (lincRNA) transcripts. We also synthesized a library consisting of 70,290 guides targeting all human RefSeq coding isoforms to screen for genes that, upon activation, confer resistance to a BRAF inhibitor. The top hits included genes previously shown to be able to confer resistance, and novel candidates were validated using individual sgRNA and complementary DNA overexpression. A gene expression signature based on the top screening hits correlated with markers of BRAF inhibitor resistance in cell lines and patient-derived samples. These results collectively demonstrate the potential of Cas9-based activators as a powerful genetic perturbation technology.

Achieving systematic, genome-scale perturbations within intact biological systems is important for elucidating gene function and epigenetic regulation. Genetic perturbations can be broadly classified as either loss-of-function (LOF) or gain-of-function (GOF) on the basis of their mode of action. To date, various genome-scale LOF screening methods have been developed, including approaches employing RNA interference^{1,2} and the RNA-guided endonuclease Cas9 from the microbial CRISPR (clustered regularly interspaced short palindromic repeat) adaptive immune system^{3,4}. Genome-scale GOF screening approaches have largely remained limited to the use of cDNA library overexpression systems. However, it is difficult to capture the complexity of transcript isoform variance using these libraries, and large cDNA sequences are often difficult to clone into size-limited viral expression vectors. The cost and complexity of synthesizing and using pooled cDNA libraries have also limited their use. Novel technologies that overcome such limitations would enable systematic, genome-scale GOF perturbations at endogenous loci.

Programmable DNA-binding proteins have emerged as an exciting platform for engineering synthetic transcription factors for modulating endogenous gene expression^{5–11}. Among the established custom DNA-binding domains, Cas9 is most easily scaled to facilitate genome-scale perturbations^{3,4} owing to its simplicity of programming relative to zinc finger proteins and transcription activator-like effectors (TALEs). Cas9 nuclease can be converted into an RNA-guided transcription activator (dCas9–activator) via inactivation of its two catalytic domains^{12,13} and fusion to transcription activation domains. These dCas9–activator fusions targeted to the promoter region of endogenous genes can then modulate gene expression^{7–11}. Although the current generation of dCas9-based transcription activators is able to achieve upregulation of some endogenous loci, the magnitude of transcriptional upregulation achieved by individual single-guide RNAs (sgRNAs)¹² typically ranges from low

to ineffective^{8–11}. Tiling a given promoter region with several sgRNAs can produce more robust transcriptional activation^{9–11}, but this requirement presents enormous challenges for scalability, and in particular for establishing pooled, genome-wide GOF screens.

To improve and expand the applications of Cas9, we recently undertook crystallographic studies to elucidate the atomic structure of the Cas9–sgRNA–target DNA tertiary complex¹⁴, thus enabling rational engineering of Cas9 and sgRNA. Here we report a series of structure-guided engineering efforts to create a potent transcription activation complex capable of mediating robust upregulation with a single sgRNA. Using this new activation system, we demonstrate activation of endogenous genes as well as non-coding RNAs, elucidate design rules for effective sgRNA target sites, and establish and apply genome-wide dCas9-based transcription activation screening to study drug resistance in a melanoma model. These results collectively demonstrate the broad applicability of CRISPR-based GOF screening for functional genomics research.

Structure-guided design of Cas9 complex

Transformation of the Cas9–sgRNA complex into an effective transcriptional activator requires finding optimal anchoring positions for the activation domains. Previous designs of dCas9-based transcription activators have relied on fusion of transactivation domains to either the amino or carboxy terminus of the dCas9 protein. To explore whether alternate anchoring positions would improve performance, we examined our previously determined crystal structure of the *Streptococcus pyogenes* dCas9(D10A/H840A) in complex with a single-guide RNA (sgRNA) and complementary target DNA¹⁴. We observed that the tetraloop and stem loop 2 of the sgRNA protrude outside of the Cas9–sgRNA ribonucleoprotein complex, with the distal 4 base pairs (bp) of each stem completely free of interactions with Cas9 amino acid side chains (Extended Data Fig. 1a). On the basis of these observations, along with

¹Broad Institute of MIT and Harvard, 75 Ames Street, Cambridge, Massachusetts 02142, USA. ²McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ³Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ⁴Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ⁵Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁶Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 2-11-16 Yayoi Bunkyo, Tokyo 113-0032, Japan. ⁷JST, PRESTO 2-11-16 Yayoi Bunkyo, Tokyo 113-0032, Japan.

*These authors contributed equally to this work.

functional data demonstrating that substitutions and deletions in the tetraloop and stem loop 2 regions of the sgRNA sequence do not affect Cas9 catalytic function¹⁴ (Fig. 1a), we reasoned that the tetraloop and stem loop 2 could tolerate the addition of protein-interacting RNA aptamers to facilitate the recruitment of effector domains to the Cas9 complex (Fig. 1b).

We selected a minimal hairpin aptamer, which selectively binds dimerized MS2 bacteriophage coat proteins¹⁵, and appended it to the sgRNA tetraloop and stem loop 2 (Extended Data Fig. 1b). We next tested whether MS2-mediated recruitment of VP64 to the tetraloop and stem loop 2 could mediate transcriptional upregulation more efficiently than a dCas9–VP64 fusion. As predicted, aptamer-mediated recruitment of MS2–VP64 to either tetraloop (sgRNA1.1) or stem loop 2 (sgRNA1.2)

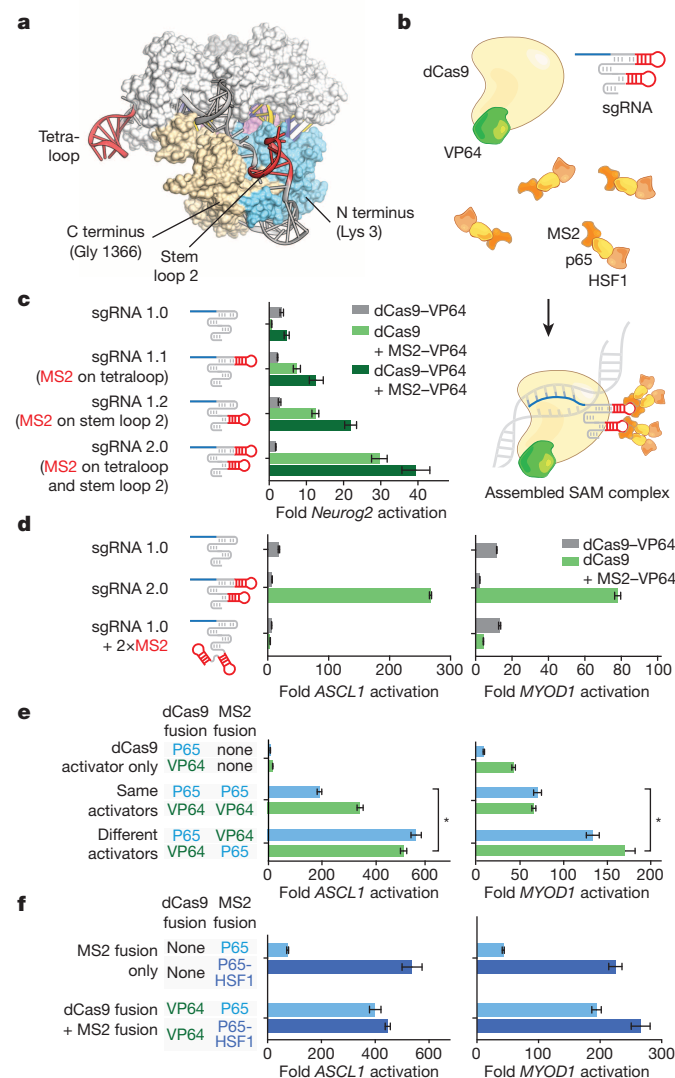


Figure 1 | Structure-guided design and optimization of an RNA-guided transcription activation complex. **a**, A crystal structure of the Cas9–sgRNA–target DNA ternary complex (PDB ID: 4OO8)¹⁴ reveals that the sgRNA tetraloop and stem loop 2 are exposed. **b**, Schematic of the three-component SAM system. **c**, Design and optimization of sgRNA scaffolds for optimal recruitment of MS2–VP64 transactivators in Neuro-2a cells. **d**, MS2 stem loop placement within the sgRNA significantly affects transcription activation efficiency. **e**, Combinations of different activation domains act in synergy to enhance the level of transcription activation. **f**, Addition of the HSF1 transactivation domain to MS2–p65 further increases the efficiency of transcription activation. Experiments for **d–f** were performed in 293FT cells. All values are mean \pm s.e.m. with $n = 3$ biological replicates. * $P < 0.05$ based on Student's *t*-test.

mediated three- and fivefold higher levels of *Neurog2* upregulation than a dCas9–VP64 fusion (sgRNA 1.0), respectively. Recruitment of VP64 to both positions (sgRNA 2.0) resulted in an additive effect, leading to a 12-fold increase over dCas9–VP64 (sgRNA 1.0). Combining sgRNA 2.0 with dCas9–VP64 instead of dCas9 provided an additional 1.3-fold increase in *Neurog2* upregulation (Fig. 1c). We further compared sgRNA 2.0 to a sgRNA previously described bearing two MS2-binding stem loops at the 3' end (sgRNA + 2 \times MS2)¹¹ and found that sgRNA 2.0 drove 14- and 8.5-fold higher levels of transcription activation than sgRNA + 2 \times MS2 for *ASCL1* and *MYOD1*, respectively (Fig. 1d). This difference could be due to either improved positioning of MS2 stem loops or to dCas9 protection of internal MS2 stem loops from exonuclease degradation.

To further improve the potency of Cas9-mediated gene activation, we considered how transcriptional activation is achieved in natural contexts, where endogenous transcription factors generally act in synergy with co-factors¹⁶. We thus hypothesized that combining VP64 with additional, distinct activation domains could improve activation efficiency. We chose the NF- κ B trans-activating subunit p65 that, while sharing some common co-factors with VP64, recruits a distinct subset of transcription factors and chromatin remodelling complexes. For example, p65 has been shown to recruit AP-1, ATF/CREB and SP1 (ref. 17), whereas VP64 recruits PC4 (ref. 18), CBP/p300 (ref. 19), and the SWI/SNF complex²⁰.

We then varied the effector domain fused to dCas9 or MS2. Hetero-effector pairing of dCas9 and MS2 fusion proteins (for example, dCas9–VP64 paired with MS2–p65 or dCas9–p65 with MS2–VP64) provided over 2.5-fold higher transcription activation for both *ASCL1* and *MYOD1* than homo-effector pairing (for example, dCas9–VP64 paired with MS2–VP64 or dCas9–p65 with MS2–p65) (Fig. 1e). We further explored this concept of domain synergy by introducing the activation domain from human heat-shock factor 1 (HSF1)²¹ as a third activation domain, and found that an MS2–p65–HSF1 fusion protein further improved transcriptional activation of *ASCL1* (12%) and *MYOD1* (37%) (Fig. 1f). Additional modifications to the sgRNA as well as Cas9 protein, including varying the nuclear localization signal (NLS), provided only minor improvements (Extended Data Fig. 1c–e). On the basis of these collective results, we concluded that the combination of sgRNA 2.0, NLS–dCas9–VP64 and MS2–p65–HSF1 comprises the most effective transcription activation system, and designated it synergistic activation mediator (SAM). For simplicity, we will refer to sgRNA 2.0 as sgRNA in subsequent discussions, unless noted otherwise.

Design rules for efficient sgRNAs

To evaluate thoroughly the effectiveness of SAM for activating endogenous gene transcription, we chose 12 genes that were previously found by several groups to be difficult to activate using dCas9–VP64 and individual sgRNA 1.0 guides^{8,10,11}. For each gene, we selected 8 sgRNA target sites spread across the proximal promoter between $-1,000$ bp and the $+1$ transcription start site (TSS). For 9 out of 12 genes, the maximum level of activation achieved using dCas9–VP64 with any of the 8 sgRNA 1.0 guides was lower than twofold, while the remaining three genes (*ZFP42*, *KLF4* and *IL1B*) were maximally activated between two- and fivefold (Fig. 2a). In contrast, SAM stimulated transcription at least twofold for all genes and more than 15-fold for 8 out of 12 genes. SAM performed consistently better than sgRNA 1.0 + dCas9–VP64 for all 96 guides, with a median gain of 105-fold greater upregulation across all 12 genes (activation by SAM divided by activation by sgRNA 1.0 + dCas9–VP64).

Previous studies have demonstrated that the poor activation efficiency of single sgRNAs can be overcome by combining dCas9–VP64 with a pool of sgRNAs tiling the proximal promoter region of the target gene^{9–11}. Therefore we compared the single sgRNA activation efficiency of SAM against dCas9–VP64 combined with a pool of 8 sgRNA 1.0 guides, all targeting the same gene. For 10 out of 12 genes, SAM with a single sgRNA performed more robustly than dCas9–VP64 with pools

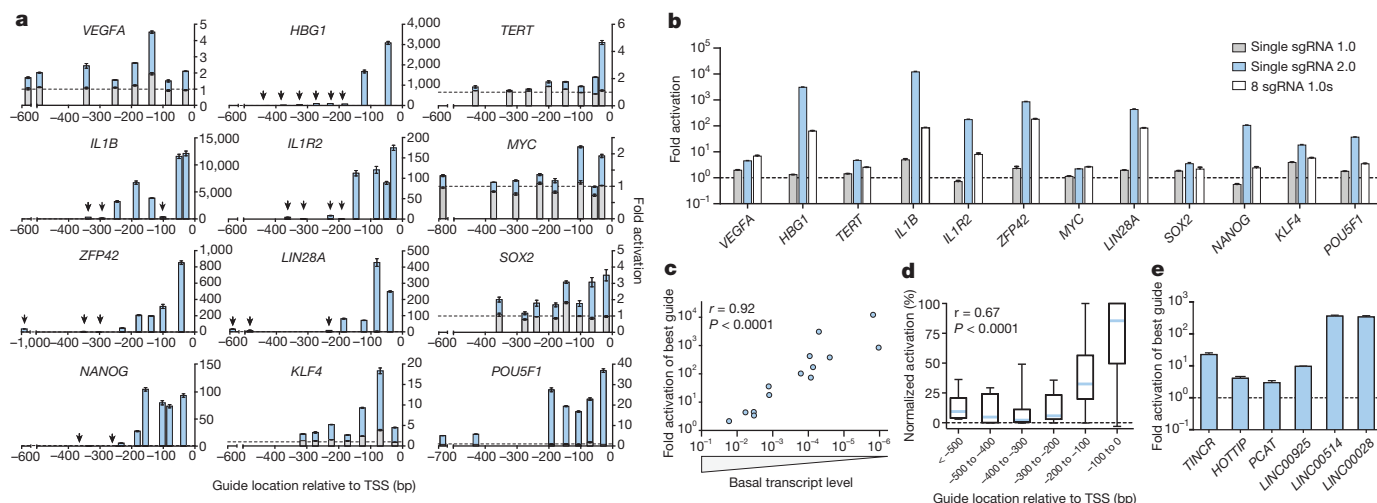


Figure 2 | Characterization of SAM-mediated gene and lincRNA activation and derivation of selection rules for efficient sgRNAs. **a**, Fold activation of 12 different genes plotted against the sgRNA location. sgRNA 1.0 with dCas9–VP64 (grey), sgRNA 2.0 with dCas9–VP64 and MS2–p65–HSF1 (blue). **b**, Comparison of activation efficiency of 12 target genes: dCas9–VP64 and a single sgRNA 1.0; dCas9–VP64 with a single sgRNA 2.0 and MS2–p65–HSF1, and dCas9–VP64 with a mixture of 8 sgRNA 1.0s. **c**, Efficiency of target gene

activation as a function of baseline expression levels. **d**, Correlation of gene activation efficiency with sgRNA targeting position. Activation efficiency of each sgRNA for the same target gene is normalized against the highest-activating sgRNA. **e**, Fold activation of six lincRNA transcripts by SAM (best sgRNA out of 8 tested). All experiments were performed in 293FT cells. All values are mean \pm s.e.m. with $n = 3$ biological replicates.

of 8 sgRNA 1.0 guides (Fig. 2b). Additionally, inclusion of a third activation domain, MS2–p65–HSF1 or MS2–p65–MyoD1, outperformed MS2–p65 alone (Extended Data Fig. 2a).

Next, we sought to determine factors that contribute to inter- and intragenic variability of activation efficiency by different sgRNAs. For inter-gene variability, differences in activation magnitudes could be due to epigenetic factors and/or variation in basal transcription levels. We were thus interested in correlating basal transcription with the level of transcription activation achieved using SAM. Using the relative transcriptional levels of target genes in control samples, we observed a highly significant correlation between the inverse of basal transcript level and the fold upregulation achieved using SAM (Fig. 2c; $r = 0.94$, $P < 0.0001$). This suggests that the basal expression level of each gene largely determines the level of activation.

To study the intragenic variability of SAM activity, we aggregated the activation data for all 96 guides and found the distance between the guide RNA target site and the TSS to be the strongest predictor of activation efficiency (Fig. 2d; $r = 0.67$, $P < 0.0001$). For all genes, the highest levels of activation were consistently achieved by targeting within the -200 bp to $+1$ bp window. This simple design guideline can inform the selection of efficient sgRNAs for gene activation.

We also sought to test whether SAM is able to activate non-coding elements in addition to protein-coding genes. We chose a diverse set of 6 lincRNAs and found that SAM mediated significant upregulation of each target (Fig. 2e), with MS2–p65–HSF1 or MS2–p65–MyoD1 leading to the highest levels of activation for each lincRNA ($P < 0.01$) (Extended Data Fig. 3). We also examined the effect of the most potent sgRNA for each lincRNA on the transcription of the nearest coding gene. Of all sgRNAs tested, only the sgRNA targeting *HOTTIP*—the only sgRNA located within 500 bp of the neighbouring gene's TSS—led to significant activation of its neighbour (Extended Data Fig. 2b).

Multiplex gene activation

The ability to simultaneously modulate gene expression at multiple loci would allow for a better understanding of complex genetic and regulatory networks. Using sets of two to ten sgRNAs, we observed successful activation of all target genes (>2 -fold) within all sgRNA combinations (Fig. 3a, b and Extended Data Fig. 4). As expected, most genes (excluding *IL1R2*) exhibited a decrease in the amount of upregulation achieved

when concurrently targeted with 9 other genes. Interestingly, the relative activation levels of each gene changed between multiplex activation and single-gene activation experiments (Fig. 3a, b).

We asked if reduced activation of targets during multiplexing was due to the reduced amounts of sgRNA or SAM protein components. Surprisingly, diluting the sgRNA expression plasmid by tenfold in single-gene activation experiments did not reduce activation for all genes (Fig. 3c). We found that genes whose levels of activation are reduced upon sgRNA dilution also exhibited dampened levels of activation when multiplexed (Fig. 3d; $r = 0.94$, $P < 0.001$). In contrast, the activation efficiency of SAM was generally unperturbed by dilution of its protein components (dCas9–VP64 and MS2–p65–HSF1) (Extended Data Fig. 5). Activation efficiency remained stable particularly when all three components were diluted, retaining on average 90% activation efficiency across a 50-fold dilution range (Extended Data Fig. 5). This finding

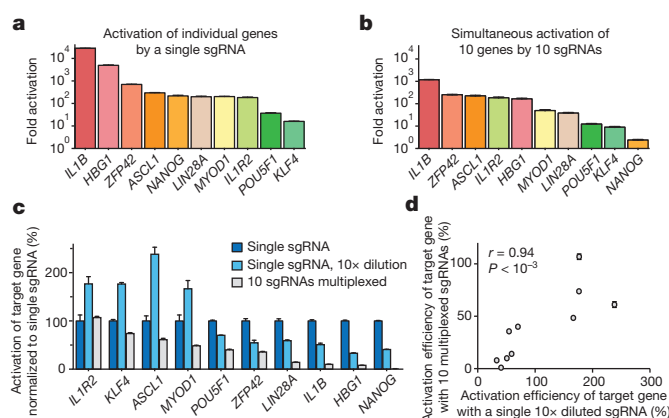


Figure 3 | Simultaneous activation of endogenous genes using multiplexed sgRNA expression. **a**, Activation of individual genes by single sgRNAs with dCas9–VP64 and MS2–p65–HSF1. **b**, Simultaneous activation of the same ten genes using a mixture of ten sgRNAs, each targeting a different gene. **c**, Effect of sgRNA dilution on gene activation efficiency. **d**, Correlation between the activation efficiency of a single tenfold diluted sgRNA and that of the same sgRNA delivered within a mixture of ten different-gene targeting sgRNAs. All values are mean \pm s.e.m. with $n = 3$ biological replicates.

was particularly promising for genome-scale pooled screening applications, which rely on single-copy lentiviral integration.

Specificity of SAM-mediated activation

An important consideration for SAM use is its targeting specificity. Recent analysis of genome-wide dCas9-binding revealed significant concentration-dependent off-target binding²², yet its effect on the specificity of transcription modulation remains unclear. To assess SAM specificity, we chose *HBG1/2* as our target gene, reasoning that globin genes would have few downstream targets that could confound our specificity analysis. We found that SAM specifically activated both *HBG1* and *HBG2* isoforms ($P < 0.05$, t -test after 0.01 false discovery rate (FDR) correction), which share the same TSS (Fig. 4 and Extended Data Fig. 6). We also tested two additional non-targeting sgRNAs with guide sequences that do not share perfect homology with the human genome. For all sgRNAs, we found only two additional genes, *S100A1* and *CYB5R2*, to be differentially expressed ($P < 0.05$, t -test after 0.01 FDR correction for multiple hypothesis testing) compared with green fluorescent protein (GFP)-expressing control (Extended Data Fig. 6) for both non-targeting guides. These results suggest that SAM-mediated gene activation is specific with minimal off-target activity.

Genome-scale gene activation screen

The ability to activate target genes using individual sgRNAs greatly facilitates the development of pooled, genome-scale transcriptional activation screening. To develop a SAM-based screening system, we generated lentiviral expression vectors that are able to drive robust transcription activation at low multiplicity of infection (MOI) (Extended Data Fig. 7a, b). Using this lentiviral system, we generated a genome-scale sgRNA library consisting of 70,290 guides, targeting every coding isoform from the RefSeq database (23,430 isoforms). For each gene, three sgRNAs were chosen to target sites within 200 bp upstream of the TSS, which was previously determined to provide more efficient activation (Fig. 2d and Fig. 5a).

Previously we applied genome-scale CRISPR knockout (GeCKO) screening³ in A375 (BRAF(V600E)) melanoma cells to identify LOF mutations capable of mediating resistance against the BRAF inhibitor PLX-4720. Here we sought to use the new SAM sgRNA library to identify

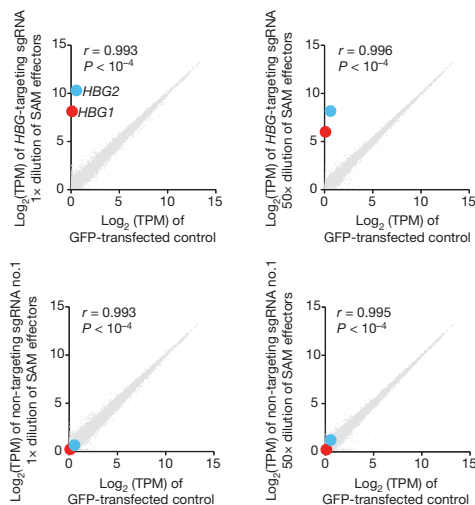


Figure 4 | Evaluation of SAM specificity. Expression levels in \log_2 (transcripts per million (TPM)) values of all detected genes in RNA-seq libraries of GFP-transfected controls (x axis of all graphs) compared to (from left to right): SAM targeting *HBG1/2* genes in 1 \times dilution and 50 \times dilution, non-targeting control sgRNAs in 1 \times dilution and 50 \times dilution (y axis). Marked are the two most statistically significant differentially expressed genes (t -test q value < 0.05 with FDR correction): Red, *HBG1*; blue, *HBG2*. The average from $n = 3$ biological replicates is shown.

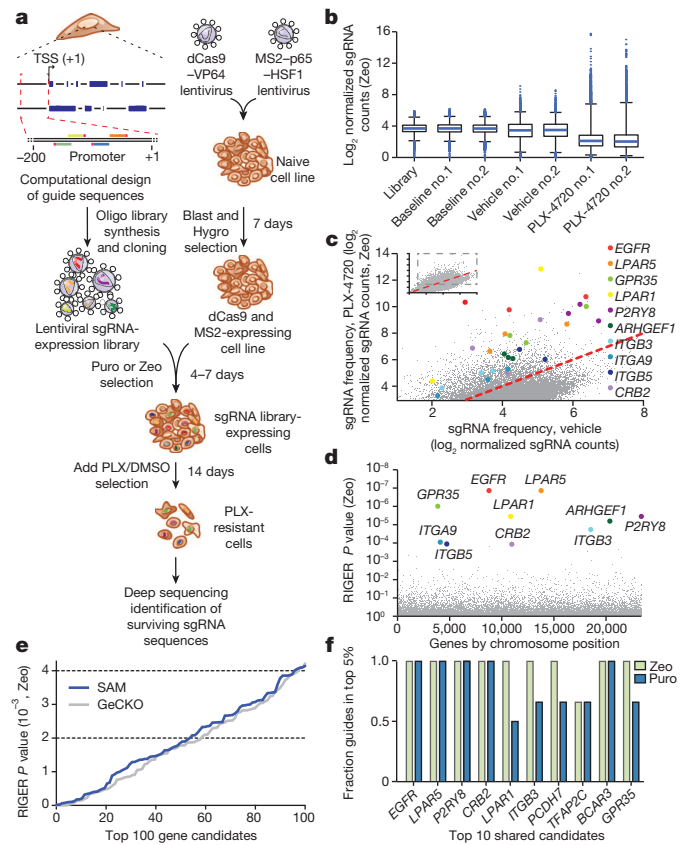


Figure 5 | Genome-scale gene activation screening identifies mediators of BRAF inhibitor resistance. a, Flow-chart of transcription activation screening using SAM. Blast, blasticidin; Hygro, hygromycin; Puro, puromycin; PLX, PLX-4720; Zeo, zeocin. b, Box-plot showing the distribution of sgRNA frequencies post lentiviral transduction for baseline (day 3), vehicle (day 21), and PLX-4720 (day 21) conditions from $n = 2$ infection replicates. c, Scatterplot showing enrichment of specific sgRNAs after PLX-4720 treatment. d, Identification of top candidate genes using the RIGER P value analysis based on the average of both infection replicates. e, Comparison of RIGER P values for the top 100 hits from SAM and GeCKO³ PLX-4720 resistance screens. f, Consistency of sgRNAs for top screening hits. Fraction of unique sgRNAs targeting each gene that are in the top 5% of all sgRNAs is plotted.

a complementary set of GOF changes that can confer BRAF inhibitor resistance (Fig. 5a).

We found that at 14 days post drug treatment, the sgRNA distribution was significantly different between cells treated with PLX-4720 and with vehicle, with the majority of sgRNAs exhibiting a reduced representation and a small set of guides showing high enrichment in PLX-4720-treated cells (Fig. 5b and Extended Data Fig. 7c). For a number of gene targets, multiple sgRNAs targeting the same gene were enriched in PLX-4720-treated cells (Fig. 5c) and the 10 most significant hits were distributed throughout the genome (Fig. 5d and Extended Data Fig. 7d). The significance of the P values of our top 100 hits determined by RNAi gene enrichment ranking (RIGER) (Supplementary Tables 1 and 2) was comparable to those observed for GeCKO screening³ (Fig. 5e). In addition, for the top 10 shared hits between two independent screens (zeocin and puromycin selection for sgRNA expression), the fraction of effectively enriched guides per gene (present in the top 5% of all guides) was very high with 97% for zeocin and 81% for puromycin ($89\% \pm 10.7\%$ overall, compared to $78\% \pm 27\%$ for the top 10 GECKO hits, Fig. 5f and Extended Data Fig. 7e).

Our screen results highlight a number of gene candidates that both confirm known PLX-4720 resistance pathways and suggest new mechanisms (Extended Data Fig. 7f). First, reactivation of the ERK pathway

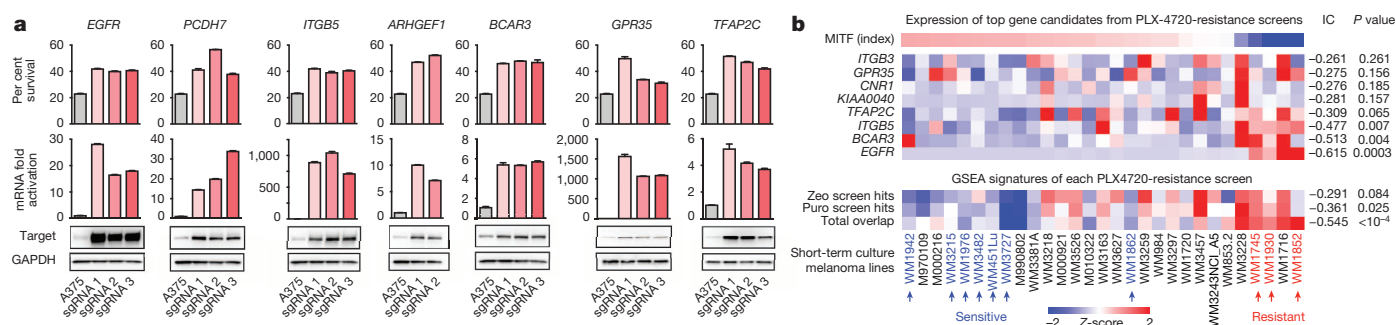


Figure 6 | Validation of top hits from genome-scale gene activation screen for PLX-4720 resistance mediators. **a**, Comparison of PLX-4720 resistance, transcription activation and protein upregulation in A375 cells for top screening hits. **b**, Expression levels of top hits and screen signatures are elevated in the resistant state of short-term BRAF(V600) melanoma cultures (see

Methods for signature generation). The subset of samples which were previously tested for PLX-4720 sensitivity and resistance are indicated by blue and red arrows, respectively³⁹. IC, information coefficient. All values are mean \pm s.e.m. with $n = 4$ biological replicates.

is one of the main known resistance mechanisms^{23,24}, and two of our screening hits, *BCAR3* and *EGFR*, probably modulate downstream and upstream nodes of this pathway, respectively^{25,26}. *EGFR* has been previously validated as a mediator of resistance to PLX-4720 through PI3K-AKT, in addition to ERK^{26,27}. These two pathways are thought to be alternative routes of PLX-4720 resistance^{24,28,29}. Furthermore, four out of the top 10 hits from our screen belong to the family of G-protein-coupled receptors (GPCRs: *GPR35*, *LPAR1*, *LPAR5* and *P2RY8*), which emerged as the top-ranked protein class conferring resistance to multiple MAP kinase inhibitors in melanoma cells in a recent screen using cDNA overexpression³⁰. GPCRs signal through multiple downstream pathways including ERK and AKT as well as cAMP-PKA^{31,32}. The final class of protein candidates from our screen belongs to the ITG receptor family, which is thought to interact with RTK and activate both ERK and AKT pathways^{33,34}.

To verify the results from the PLX-4720 resistance screen, we validated each of the top 13 genes. All sgRNAs from the screen that targeted these 13 genes conferred PLX-4720 resistance when individually expressed in A375 along with SAM (Fig. 6a and Extended Data Fig. 8a). We also verified that SAM was able to facilitate robust increase in target transcript (Fig. 6a and Extended Data Fig. 8b) and protein levels (Fig. 6a). Since 5 of our top candidates from the pooled SAM screen overlapped with hits from a previously conducted arrayed cDNA screen³⁰ (Extended Data Fig. 8c), we compared the relative efficacy of cDNA overexpression with SAM-mediated transcription activation. Interestingly, for these 5 targets, SAM led to at least similar levels of PLX-4720 resistance when compared with corresponding cDNA overexpression conditions (Extended Data Fig. 8a), despite cDNA leading to higher transcript levels (Extended Data Fig. 8d). Furthermore, we found that, for most genes, the levels of PLX-4720 resistance mediated by all three sgRNAs were comparable (Extended Data Fig. 8e).

In addition to validating our top screening hits through individual sgRNA or cDNA overexpression, we analysed the expression profile of our screening hits using four different data sets (CCLE^{35,37}, TCGA: <https://tcga-data.nci.nih.gov/tcga/>, short-term melanoma cultures³⁶, and pre/post treatment patient samples³⁸). As shown previously³⁹, a distinct transcriptional state defines BRAF-inhibition sensitive and resistant states as described by activation of endogenous MITF/associated markers (for example, PMEL) and NF- κ B-pathway activity/associated markers (for example, AXL), respectively (Fig. 6b and Extended Data Fig. 9b). Based on short-term melanoma data^{36,39}, we found that the expression of our top screening hits was significantly increased in the resistant state. Correspondingly, a gene expression signature based on the top screening hits (see Methods) correlated with a BRAF-inhibitor resistance state as defined previously³⁹ (Fig. 6b; total overlap, $P < 0.0001$). Further analysis performed using the CCLE, TCGA and pre/post treatment data set also revealed similar correlations (Extended Data Fig. 9).

Discussion

In summary, we have taken a structure-guided approach to design a dCas9-based transcription activation system for achieving robust, single sgRNA-mediated gene upregulation. By engineering the sgRNA to incorporate protein-interacting aptamers, we assembled a synthetic transcription activation complex consisting of multiple distinct effector domains modelled after natural transcription activation processes. Here we have shown that the SAM system is robust, specific, and can facilitate genome-scale gain-of-function screening when combined with a compact pooled sgRNA library. Our SAM-mediated screens exhibited a high degree of consistency and validation, with $>80\%$ effectively enriched guides per gene hit, and 100% validation of the top 10 hits.

Future engineering of the Cas9 complex based on structural information^{14,40} will further expand the Cas9 toolbox⁴¹. Additional developments of the SAM system may be able to take advantage of the modularity and customizability of the sgRNA scaffold to establish a series of sgRNA scaffolds bearing different aptamers for recruiting distinct types of effectors in an orthogonal manner. For instance, replacement of the MS2 stem loops with PP7-interacting stem loops may be used to recruit repressive elements, potentially enabling multiplexed bidirectional transcriptional control.

Although we have taken initial steps towards defining selection rules for potent sgRNAs, future studies will reveal additional selection criteria that are critical for guide efficacy, such as sequence-intrinsic properties (Extended Data Fig. 10a–d). Applications of dCas9-based transcription modulators in positive and negative selection screens (Extended Data Fig. 10e, f)⁴² will enable the dissection of many types of genetic elements, ranging from protein-coding genes to non-coding lincRNA elements. Furthermore, combining wild-type Cas9-mediated genome modifications with SAM-mediated recruitment of epigenetic modifiers will constitute powerful approaches for studying genome organization and regulation in diverse biological processes.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 September; accepted 3 December 2014.

Published online 10 December 2014.

- Berns, K. *et al.* A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **428**, 431–437 (2004).
- Boutros, M. *et al.* Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* **303**, 832–835 (2004).
- Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014).
- Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014).
- Beerli, R. R., Segal, D. J., Dreier, B. & Barbas, C. F. III. Toward controlling gene expression at will: specific regulation of the *erbB-2/HER-2* promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl Acad. Sci. USA* **95**, 14628–14633 (1998).

6. Zhang, F. *et al.* Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nature Biotechnol.* **29**, 149–153 (2011).
7. Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013).
8. Konermann, S. *et al.* Optical control of mammalian endogenous transcription and epigenetic states. *Nature* **500**, 472–476 (2013).
9. Maeder, M. L. *et al.* CRISPR RNA-guided activation of endogenous human genes. *Nature Methods* **10**, 977–979 (2013).
10. Perez-Pinera, P. *et al.* RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nature Methods* **10**, 973–976 (2013).
11. Mali, P. *et al.* CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnol.* **31**, 833–838 (2013).
12. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
13. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl Acad. Sci. USA* **109**, E2579–E2586 (2012).
14. Nishimasu, H. *et al.* Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935–949 (2014).
15. Peabody, D. S. The RNA binding site of bacteriophage MS2 coat protein. *EMBO J.* **12**, 595–600 (1993).
16. Lemon, B. & Tjian, R. Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev.* **14**, 2551–2569 (2000).
17. van Essen, D., Engist, B., Natoli, G. & Sacconi, S. Two modes of transcriptional activation at native promoters by NF- κ B p65. *PLoS Biol.* **7**, e73 (2009).
18. Kretzschmar, M., Kaiser, K., Lottspeich, F. & Meisterernst, M. A novel mediator of class II gene transcription with homology to viral immediate-early transcriptional regulators. *Cell* **78**, 525–534 (1994).
19. Ikeda, K., Stuehler, T. & Meisterernst, M. The H1 and H2 regions of the activation domain of herpes simplex virion protein 16 stimulate transcription through distinct molecular mechanisms. *Genes Cells* **7**, 49–58 (2002).
20. Neely, K. E. *et al.* Activation domain-mediated targeting of the SWI/SNF complex to promoters stimulates transcription from nucleosome arrays. *Mol. Cell* **4**, 649–655 (1999).
21. Marinho, H. S., Real, C., Cyrne, L., Soares, H. & Antunes, F. Hydrogen peroxide sensing, signaling and regulation of transcription factors. *Redox Biol.* **2**, 535–562 (2014).
22. Wu, X. *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nature Biotechnol.* **32**, 670–676 (2014).
23. Johannessen, C. M. *et al.* COT drives resistance to RAF inhibition through MAP kinase pathway reactivation. *Nature* **468**, 968–972 (2010).
24. Nazarian, R. *et al.* Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation. *Nature* **468**, 973–977 (2010).
25. Musgrove, E. A. & Sutherland, R. L. Biological determinants of endocrine resistance in breast cancer. *Nature Rev. Cancer* **9**, 631–643 (2009).
26. Prahallad, A. *et al.* Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature* **483**, 100–103 (2012).
27. Corcoran, R. B. *et al.* EGFR-mediated re-activation of MAPK signaling contributes to insensitivity of BRAF mutant colorectal cancers to RAF inhibition with vemurafenib. *Cancer Discov.* **2**, 227–235 (2012).
28. Villanueva, J. *et al.* Acquired resistance to BRAF inhibitors mediated by a RAF kinase switch in melanoma can be overcome by cotargeting MEK and IGF-1R/PI3K. *Cancer Cell* **18**, 683–695 (2010).
29. Shi, H., Kong, X., Ribas, A. & Lo, R. S. Combinatorial treatments that overcome PDGFR β -driven resistance of melanoma cells to ^{V600E}B-RAF inhibition. *Cancer Res.* **71**, 5067–5074 (2011).
30. Johannessen, C. M. *et al.* A melanocyte lineage program confers resistance to MAP kinase pathway inhibition. *Nature* **504**, 138–142 (2013).
31. Dorsam, R. T. & Gutkind, J. S. G-protein-coupled receptors and cancer. *Nature Rev. Cancer* **7**, 79–94 (2007).
32. Lappano, R. & Maggiolini, M. G protein-coupled receptors: novel targets for drug discovery in cancer. *Nature Rev. Drug Discov.* **10**, 47–60 (2011).
33. Franke, T. F. PI3K/Akt: getting it right matters. *Oncogene* **27**, 6473–6488 (2008).
34. Desgrosellier, J. S. & Cheresch, D. A. Integrins in cancer: biological implications and therapeutic opportunities. *Nature Rev. Cancer* **10**, 9–22 (2010).
35. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
36. Lin, W. M. *et al.* Modeling genomic diversity and tumor dependency in malignant melanoma. *Cancer Res.* **68**, 664–673 (2008).
37. Wilks, C. *et al.* The Cancer Genomics Hub (CGHub): overcoming cancer through the power of torrential data. *Database* **2014**, bau093 (2014).
38. Rizos, H. *et al.* BRAF inhibitor resistance mechanisms in metastatic melanoma: spectrum and clinical impact. *Clin. Cancer Res.* **20**, 1965–1977 (2014).
39. Konieczkowski, D. J. *et al.* A melanoma cell state distinction influences sensitivity to MAPK pathway inhibitors. *Cancer Discov.* **4**, 816–827 (2014).
40. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569–573 (2014).
41. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
42. Gilbert, L. A. *et al.* Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* **159**, 647–661 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank S. Shehata, K. Zheng, C. Johannessen, L. Garraway, O. Shalem and members of the Zhang laboratory for assistance and helpful discussions. O.O.A. is supported by a NSF Graduate Research Fellowship, J.S.G. is supported by a D.O.E. Computational Science Graduate Fellowship, H.N. is supported by PRESTO from JST and Grant-in-Aid for Scientific Research (B) from JSPS, O.N. is supported by the CREST program and JST, and F.Z. is supported by the NIMH (DP1-MH100706), the NINDS (R01-NS07312401), NSF, the Keck, Searle Scholars, Klingenstein, Vallee, and Simons Foundations, and Bob Metcalfe. CRISPR reagents are available to the academic community through Addgene, and associated protocols, support forum and computational tools are available via the Zhang laboratory website (<http://www.genome-engineering.org>).

Author Contributions S.K. and F.Z. conceived the project. S.K., M.D.B., A.E.T. and F.Z. designed the experiments. S.K., M.D.B., A.E.T., C.B., P.D.H. and J.J. performed experiments and analysed data. H.N. and O.N. helped with structural interpretation. N.H. performed the RNA-seq analysis. J.S.G. performed the depletion guide efficacy analysis. O.O.A. performed the analysis of clinical data sets. S.K., A.E.T., P.D.H. and F.Z. wrote the paper with help from all authors.

Author Information All reagents described in this manuscript have been deposited with Addgene (plasmid IDs 61422–61427 for SAM component plasmid and 61597 for the human SAM guide RNA library). RNA-seq data are available at BioProject under accession number PRJNA269048. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.Z. (zhang@broadinstitute.org).

METHODS

Sequences. DNA sequences for SAM components and sgRNA scaffolds are provided in Supplementary Sequences. sgRNA target sequences for characterization and optimization of SAM are listed in Supplementary Table 4.

Transient transfection experiments. Neuro-2a cells (Sigma-Aldrich) were grown in media containing 1:1 ratio of OptiMEM (Life Technologies) to high-glucose DMEM with GlutaMAX and sodium pyruvate (Life Technologies) supplemented with 5% HyClone heat-inactivated FBS (Thermo Scientific), 1% penicillin/streptomycin (Life Technologies), and passaged at 1:5 every 2 days.

HEK293FT cells (Life Technologies) were maintained in high-glucose DMEM with GlutaMAX and sodium pyruvate (Life Technologies) supplemented with 10% heat-inactivated characterized HyClone fetal bovine serum (Thermo Scientific) and 1% penicillin/streptomycin (Life Technologies). Cells were passaged daily at a ratio 1:2 or 1:2.5. For gene activation experiments, 20,000 HEK293FT cells per well were plated in 100 μ l media in poly-D-lysine-coated 96-well plates (BD BioSciences). 24 h after plating, cells were transfected with a 1:1:1 mass ratio of sgRNA plasmid with gene-specific targeting sequence or pUC19 control plasmid:MS2-effector plasmid or pUC19:dCas9 plasmid, dCas9-effector plasmid or pUC19.

A total plasmid mass of 0.3 μ g per well was transfected using 0.6 μ l per well of Lipofectamine 2000 (Life Technologies) according to the manufacturer's instructions. Culture medium was changed 5 h after transfection. 48 h after transfection, cell lysis and reverse transcription were performed using a Cells-to-Ct kit (Life Technologies). Relative RNA expression levels were quantified by reverse transcription and quantitative PCR (qPCR) using TaqMan qPCR probes (Life Technologies, Supplementary Table 5) and Fast Advanced Master Mix (Life Technologies). qPCR was carried out in 5 μ l multiplexed reactions and 384-well format using a LightCycler 480 Instrument II. Data was analysed by the $\Delta\Delta C_t$ method: target C_t values (FAM dye) were normalized to GAPDH C_t values (VIC dye), and fold changes in target gene expression were determined by comparing to GFP-transfected experimental controls.

Lentivirus production. HEK293T cells (Life Technologies) were cultured as described above for HEK293FT cells. 1 day before transfection, cells were seeded at ~40% confluency (12 \times T225 flasks for library scale production, 1 \times T25 flask for individual guide production). Cells were transfected the next day at ~80–90% confluency. For each flask, 10 μ g of plasmid containing the vector of interest, 10 μ g of pMD2.G and 15 μ g of psPAX2 (Addgene) were transfected using 100 μ l of Lipofectamine 2000 and 200 μ l Plus Reagent (Life Technologies). 5 h after transfection the media was changed. Virus supernatant was harvested 48 h post-transfection, filtered with a 0.45- μ m PVDF filter (Millipore), aliquoted, and stored at -80°C .

Lentiviral transduction. A375 cells (Sigma) were cultured in RPMI 1640 (Life Technologies) supplemented with 10% FBS (Seradigm) and 1% penicillin/streptomycin (Life Technologies) and passaged every other day at a 1:4 ratio. Cells were transduced with lentivirus via spinfection in 12-well plates. 3×10^6 cells in 2 ml of media supplemented with 8 $\mu\text{g ml}^{-1}$ polybrene (Sigma) were added to each well, supplemented with lentiviral supernatant and centrifuged for 2 h at 1,000g. 24 h after spinfection, cells were detached with TrypLE (Life Technologies) and counted. Cells were replated at low density (7.5×10^6 cells per T225 Flask) and a selection agent was added either immediately (zeocin, blasticidin and hygromycin, all Life Technologies) or 3 h after plating (puromycin). Concentrations for selection agents were determined using a kill curve: 0.5 $\mu\text{g ml}^{-1}$ puromycin, 200 $\mu\text{g ml}^{-1}$ zeocin, 10 $\mu\text{g ml}^{-1}$ blasticidin, and 300 $\mu\text{g ml}^{-1}$ hygromycin. Media was refreshed on day 2 and cells were passaged every other day starting on day 4 after replating. The duration of selection was 4 days for puromycin and 7 days for zeocin, hygromycin and blasticidin. Lentiviral titres were determined by spinfecting cells with 6 different volumes of lentivirus ranging from 0 to 600 μ l and counting the number of surviving cells after a complete selection (3–6 days).

Design and cloning of SAM library. RefSeq coding gene isoforms with a unique TSS (total of 23,430 isoforms) were targeted with three guides each for a total library of 70,290 guides (Supplementary Table 6). Guides were designed to target the first 200 bp upstream of each TSS and subsequently filtered for GC content $>25\%$ and minimal overlap of the target sequence. After filtering, the remaining guides were scored according to predicted off-target matches as described previously⁴³, and three guides with the best off-target scores were selected. Cloning of the SAM sgRNA libraries was performed as previously described³ with a minimum representation of 100 transformed colonies per guide.

Depletion and PLX-4720 screen. A375 cells stably integrated with SAM Cas9 and effector components were transduced with SAM sgRNA libraries as described above at an MOI of 0.2, with a minimal representation of 500 transduced cells per guide. Cells were maintained at >500 cells per guide during subsequent passaging. At 7 days post infection (DPI) (complete selection, see above), cells were split into vehicle (DMSO) and PLX-4720 conditions (2 μM PLX-4720 dissolved in DMSO, Selleckchem). Cells were passaged every 2 days for a total of 14 days of drug treatment. >500 cells per guide were harvested as a baseline at 3 DPI (4 days before treatment)

and at 21 DPI (after 14 days of treatment) for gDNA extraction. Genomic DNA was extracted using the Zymo Quick-gDNA midi kit (Zymo Research). PCR of the virally integrated guides was performed on gDNA at the equivalent of >500 cells per guide in 96 parallel reactions using NEBnext High Fidelity 2X Master Mix (New England Biolabs) in a single-step reaction of 22 cycles. Primers are listed here: forward primer, AATGATACGGCGACCCAGATCTACACTCTTTCCTACACGACGCTCTCCGATCTNNNNNNN(1–10-bp stagger)GCTTTATATATCTTGTGGAAAGGACGAAACACC, 8 bp barcode indicated in italic; reverse primer, CAAGCAGAAGACGGCATAACGATNNNNNNNGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCCAAGTTGATAACGGACTAGCCTT, 8-bp index read barcode indicated in italic.

PCR products from all 96 reactions were pooled, purified using Zymo-Spin V with Reservoir (Zymo research) and gel extracted using the Zymoclean Gel DNA Recovery Kit (Zymo research). Resulting libraries were deep-sequenced on Illumina MiSeq and HiSeq platforms with a total coverage of >35 million reads passing filter per library.

NGS and screen hits analysis. NGS data were de-multiplexed using unique index reads. Guide counts (Supplementary Table 7) were determined based on perfectly matched sequencing reads only. For each condition, guide counts were normalized to the total number of counts per condition, and \log_2 counts were calculated based on these values. Ratios of counts between conditions were calculated as $\log_2((\text{count } 1 + 1)/(\text{count } 2 + 1))$ based on normalized counts.

RIGER analysis was performed using GENE-E based on the normalized day 14 \log_2 ratios (PLX-4720/DMSO) averaged over two independent infection replicates. All RIGER analysis used the Kolmogorov–Smirnov method as described previously⁴⁴, except for Fig. 6c, where the weighted average method was used in order to enable comparison to GeCKO values determined by that method.

Gene expression and pharmacological validation analysis. Gene expression data (CCLE, TCGA, short-term cultures, patient melanoma biopsies) and pharmacological data (CCLE, short-term cultures) were analysed to better understand the biological relevance of the top gene hits from the SAM screens. In the CCLE data set³⁵, gene expression data (RNA-sequencing, GCHub: <https://cghub.ucsc.edu/datasets/ccl.html>) and pharmacological data (activity area for MAPK pathway inhibitors) from BRAF(V600) mutant melanoma cell lines were used to compute the association between PLX-4720 resistance and the gene expression of each of the top hits. Additionally, gene expression signatures comprised of the top hits were generated using single-sample Gene Set Enrichment Analysis (ssGSEA)⁴⁵, and the associations between PLX-4720 resistance and these signatures were computed.

Gene expression data (Affymetrix GeneChip HT-HGU133) and PLX-4720 pharmacological data (GI₅₀, half-maximal growth inhibition concentration; only for a subset of the samples) from short term melanoma cultures (STC)³⁶ were also used for plotting the gene expression of top hits and their ssGSEA signature scores. Expression data for the STC samples were collapsed to maximum probe value per gene and pre-processed using robust spline normalization.

Gene expression (RNA-sequencing) and genotyping data were collected from 113 BRAF(V600)-mutant primary and metastatic patient tumours from The Cancer Genome Atlas (<https://tcga-data.nci.nih.gov/tcga/>) and these data were similarly used for determining the association between resistance and the expression of top hits/ssGSEA signature scores. Because pharmacological data was not available for the STCs (only a subset had PLX-4720 data) and the TCGA melanoma samples, a transcriptional state was plotted using marker genes and signatures³⁹ in order to identify samples resistant to BRAF-inhibition.

Gene expression data from 13 patients with BRAF(V600E) melanomas³⁸ was used for analysing the relationship between resistance and the expression of our top hits/ssGSEA signature scores. Because all the post-treatment tumours were resistant and not every sample had a paired on-treatment biopsy, we decided to order the samples by MITF expression in the pre-treatment samples to reflect the original PLX-4720 sensitivity state of the tumours. We then used the expression data in the post-treatment resistant tumours to plot the expression of top hits/ssGSEA signature scores. We also calculated the \log_2 -fold change between each patient's post/pre paired samples and determined the number of patients that had at least a \log_2 -fold change of 2 per top screen hit.

Single sample gene set enrichment analysis. While there was a significant association between the overexpression of some of our top individual SAM screen hits and resistance in three external cancer data sets, we sought a more robust scoring system independent of any single gene. Gene expression signatures were generated based on the set of top hits from each of the two SAM screens and for the overlap between them. Using single-sample Gene Set Enrichment analysis (ssGSEA), a score was generated for each sample that represents the enrichment of the SAM screen gene expression signature in that sample and the extent to which those genes are coordinately up- or downregulated. Additionally, signature gene sets from the Molecular Signature Database (MSigDB)⁴⁶ were used in order to fully map the

transcriptional BRAF-inhibitor resistant/sensitive states in the short-term culture and TCGA data sets as previously described³⁹.

Information coefficient for measuring associations in external data sets. To measure correlations between different features (signature scores, gene expression, or drug-resistance data) in the external cancer data sets, an information-theoretic approach (Information Coefficient; IC) was used and significance was measured using a permutation test ($n = 10,000$), as previously described³⁹. The IC was calculated between the feature used to sort the samples (columns) in each data set and each of the features plotted in the heatmap (pharmacological data, gene expression, and signature scores).

sgRNA sequence analysis. Depletion for each sgRNA was calculated as the ratio of counts (see “NGS and screen hits analysis”) between day 3 and day 21. The sgRNAs corresponding to genes with significant depletion ($P < 0.05$ by RIGER analysis) in sgRNA-Puro and sgRNA-Zeo libraries were selected for analyses. These sgRNAs were analysed for nucleotide occurrence in the sgRNA sequence, distance from TSS, and guide strand relative to transcript orientation. For each variable, the correlation and significance with the sgRNA ratio was calculated by ordinary least squares linear regression.

PLX-4720 survival assay. A375 cells stably integrated with dCas9–VP64 and MS2–p65–HSF1 were transduced with individual guides from the top screening hits of the Zeocin and Puromycin screens (13 genes total, 3 sgRNAs per gene) as well as available cDNA at an MOI of <0.2 as described above. Cells were selected for guide expression with Zeocin (Life Technologies) for 5 days and replated at low density (3×10^3 cells per well in a 96-well plate). A375 cells and A375 cells expressing dCas9–VP64 and MS2–p65–HSF1 were plated as controls. Different concentrations of PLX-4720 (2 μ M, 0.5 μ M, 0.15 μ M) or vehicle (DMSO) were added 3 h after plating. Cells were treated with PLX-4720 for 4 days before cell viability was measured using CellTiter–Glo Luminescent Cell Viability Assay (Promega). For qPCR quantification of target gene upregulation, cells were also plated at 5 DPI (3×10^4 cells per well in a 96-well plate) and harvested for mRNA 24 h after plating.

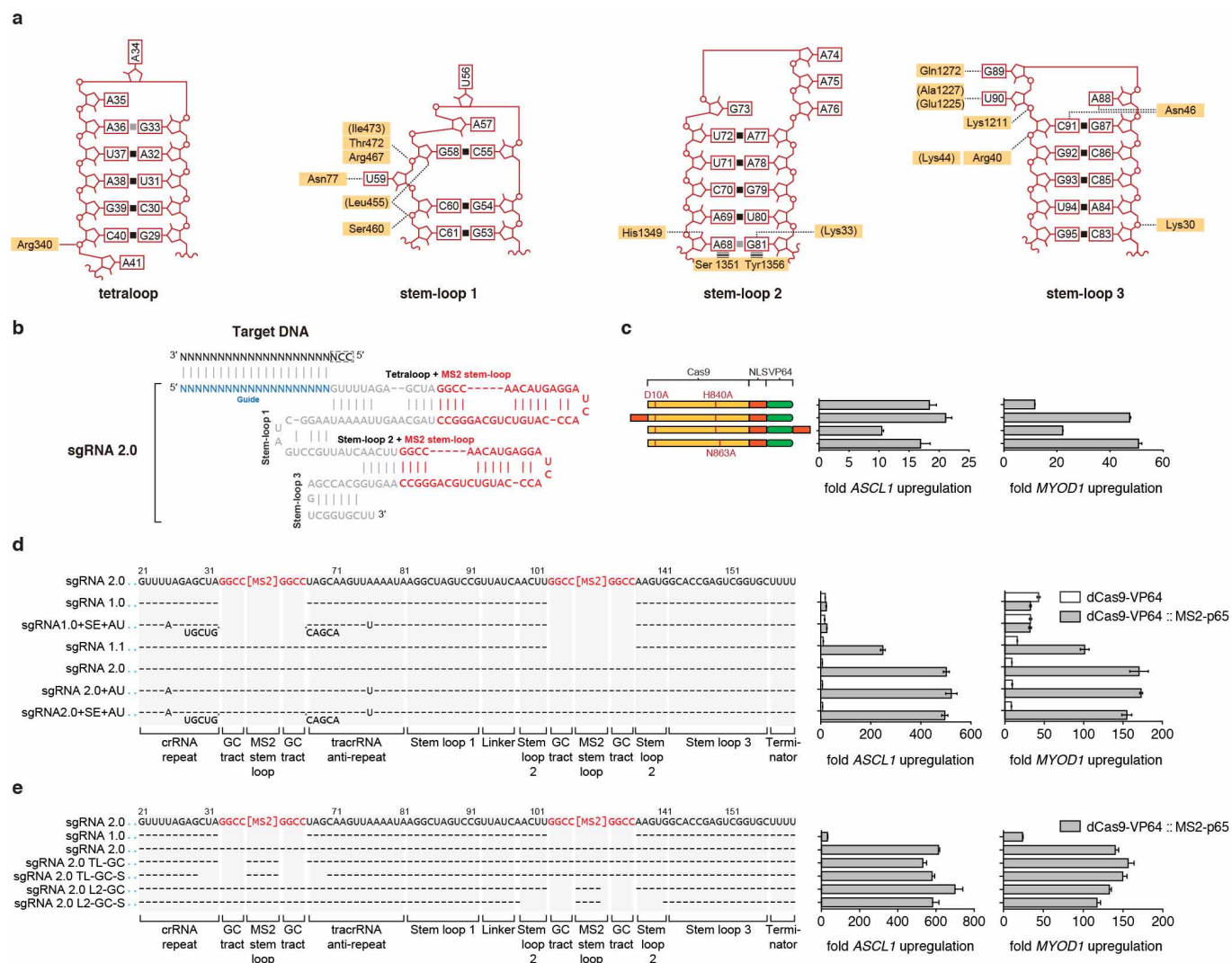
Western blot. Protein lysates were prepared with RIPA lysis buffer (Cell Signaling Technologies) containing a protease inhibitor cocktail (Roche). Samples standardized for protein with the Pierce BCA protein assay (Thermo Scientific) were boiled at 95 °C for 5 min under reducing conditions (except for GPR35 samples, which were incubated at 37 °C for 30 min). After denaturation, samples for probing proteins with lower or higher molecular weight were separated by 10–20% or 4–15% Criterion Tris–HCl gels (Bio–Rad) and electrotransferred onto a 0.2- μ m or 0.45- μ m polyvinylidene difluoride membrane (Millipore), respectively. Blots were blocked with 5% BLOT–QuickBlocker (VWR) and probed with different primary antibodies (anti–EGFR (rabbit polyclonal, SC–03, Santa Cruz Biotechnology, 1:1,000 dilution), anti–PCDH7 (rabbit polyclonal, HPA011866, Sigma–Aldrich, 1:1,000 dilution), anti–ITGB5 (rabbit polyclonal, SC–14010, Santa Cruz Biotechnology, 1:500 dilution), anti–ARHGEF1 (rabbit polyclonal, 11363–1–AP, Proteintech, 1:5,000 dilution), anti–BCAR3 (rabbit polyclonal, A301–671A, Bethyl Laboratories, 1:2,000 dilution), anti–GPR35 (rabbit polyclonal, 10007660, Cayman Chemical, 1:1,000 dilution), anti–TFAP2C (rabbit polyclonal, 2320, Cell Signaling Technology, 1:1,000 dilution), 2.5% bovine serum albumin, Sigma–Aldrich) in 2.5% BLOT–QuickBlocker (VWR) unless noted otherwise overnight at 4 °C. Blots were then incubated with secondary antibody HRP–conjugated goat anti–rabbit IgG (7074, Cell Signaling Technology,

1:1,000 dilution) and HRP–conjugated GAPDH (rabbit monoclonal, 3683, Cell Signaling Technology, 1:2,000 dilution) in 2.5% BLOT–QuickBlocker (VWR) for 1 h at room temperature. Proteins with molecular weights similar to GAPDH (GPR35 and TFAP2C) were stripped with Restore Plus Western Blot Stripping Buffer (Thermo Scientific) before probing for GAPDH. SuperSignal West Pico and Femto Chemiluminescent Substrates (Thermo Scientific) were used for detection.

RNA sequencing and data analysis. Samples harvested for RNA sequencing were prepped with TruSeq Stranded mRNA Sample Prep Kit (Illumina) and deep-sequenced on the Illumina MiSeq platform (>9 million reads per condition). Bowtie2⁴⁷ index was created based on the human hg19 UCSC genome and known gene transcriptome, and paired-end reads were aligned directly to this index using Bowtie2 with command line options “–q–phred33–quals –n 2 –e 99999999 –l 25 –I 1 –X 1000 –a –m 200 –p 4 –chunkmbs 512”. Next, RSEM v1.27⁴⁸ was run with default parameters on the alignments created by Bowtie2 to estimate expression levels. RSEM’s gene level expression estimates (tau) were multiplied by 1,000,000 to obtain transcript per million (TPM) estimates for each gene, and TPM estimates were transformed to log-space by taking $\log_2(\text{TPM} + 1)$. The normalization between libraries was tested using an MA plot (mairplot function in Matlab V2013b). Genes were considered detected if their transformed expression level was equal to or above 1 (in $\log_2(\text{TPM} + 1)$ scale). All genes detected in at least one library (out of three libraries per condition) were used to construct scatter plots comparing each of the six conditions to the control GFP condition, using the average across biological replicates with $>80\%$ alignment to the hg19 UCSC known gene transcriptome ($\log_2(\text{mean}(\text{TPM}) + 1)$ value per gene).

To find differentially expressed genes, we performed Student’s t -test on each of the six conditions against the GFP condition. The t -test was run on all genes that had expression levels above $\log_2(\text{TPM} + 1) > 2.5$ in at least two libraries. This threshold was chosen as the minimal threshold for which the number of detected genes across all libraries was constant. Only genes that were significant (P -value pass 0.01 FDR correction) and had at least 1.5-fold change were reported and visualized using a heat map.

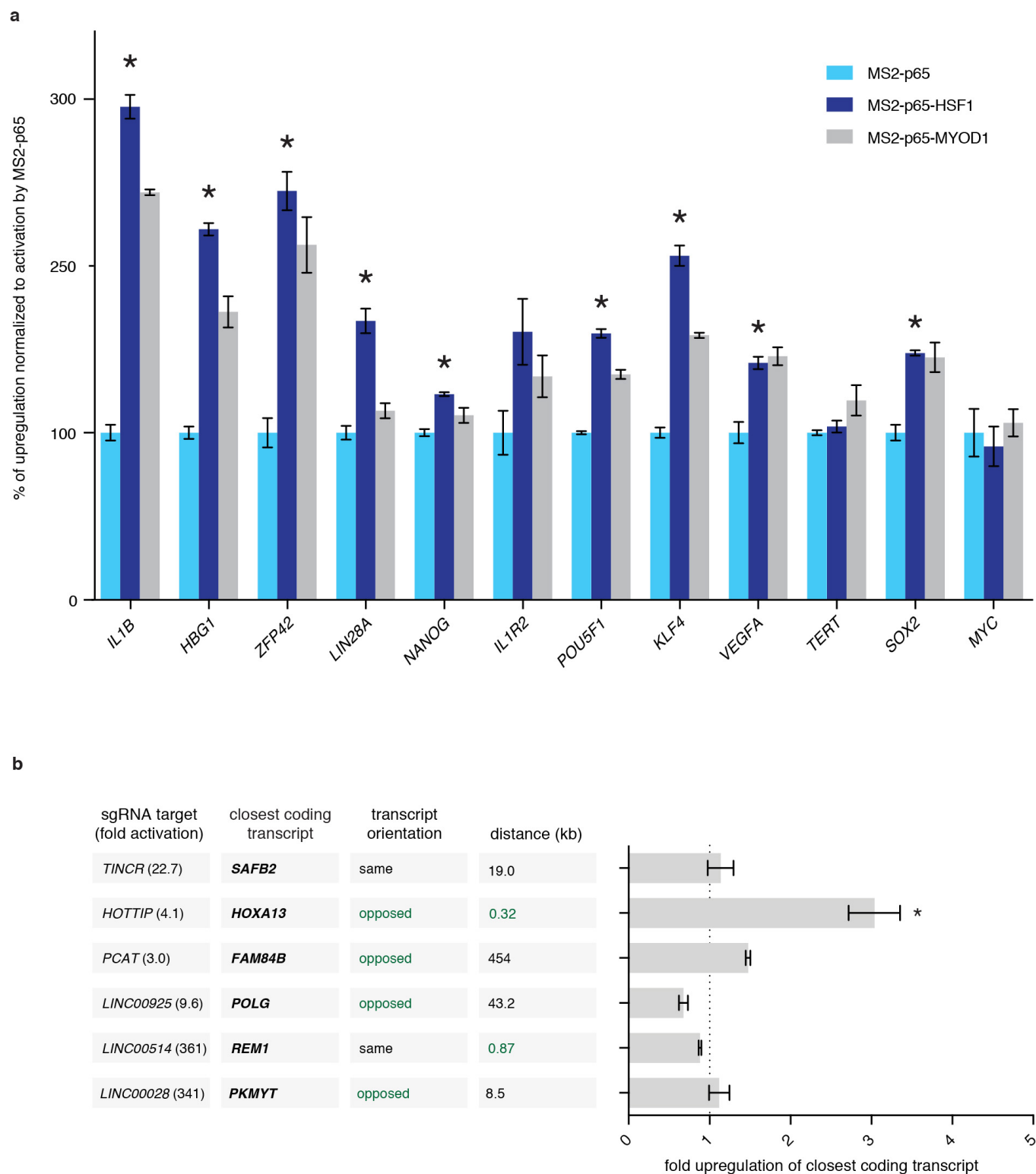
43. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnol.* **31**, 827–832 (2013).
44. Luo, B. *et al.* Highly parallel identification of essential genes in cancer cells. *Proc. Natl Acad. Sci. USA* **105**, 20380–20385 (2008).
45. Barbie, D. A. *et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **462**, 108–112 (2009).
46. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
47. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
48. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA–Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
49. Smalley, K. S. Understanding melanoma signaling networks as the basis for molecular targeted therapy. *J. Invest. Dermatol.* **130**, 28–37 (2010).
50. Wong, P. P. *et al.* Histone demethylase KDM5B collaborates with TFAP2C and Myc to repress the cell cycle inhibitor $p21^{cip}$ (CDKN1A). *Mol. Cell. Biol.* **32**, 1633–1644 (2012).
51. Hart, M. J. *et al.* Direct stimulation of the guanine nucleotide exchange activity of p115 RhoGEF by G α 13. *Science* **280**, 2112–2114 (1998).



Extended Data Figure 1 | Structure-guided engineering of Cas9 sgRNA.

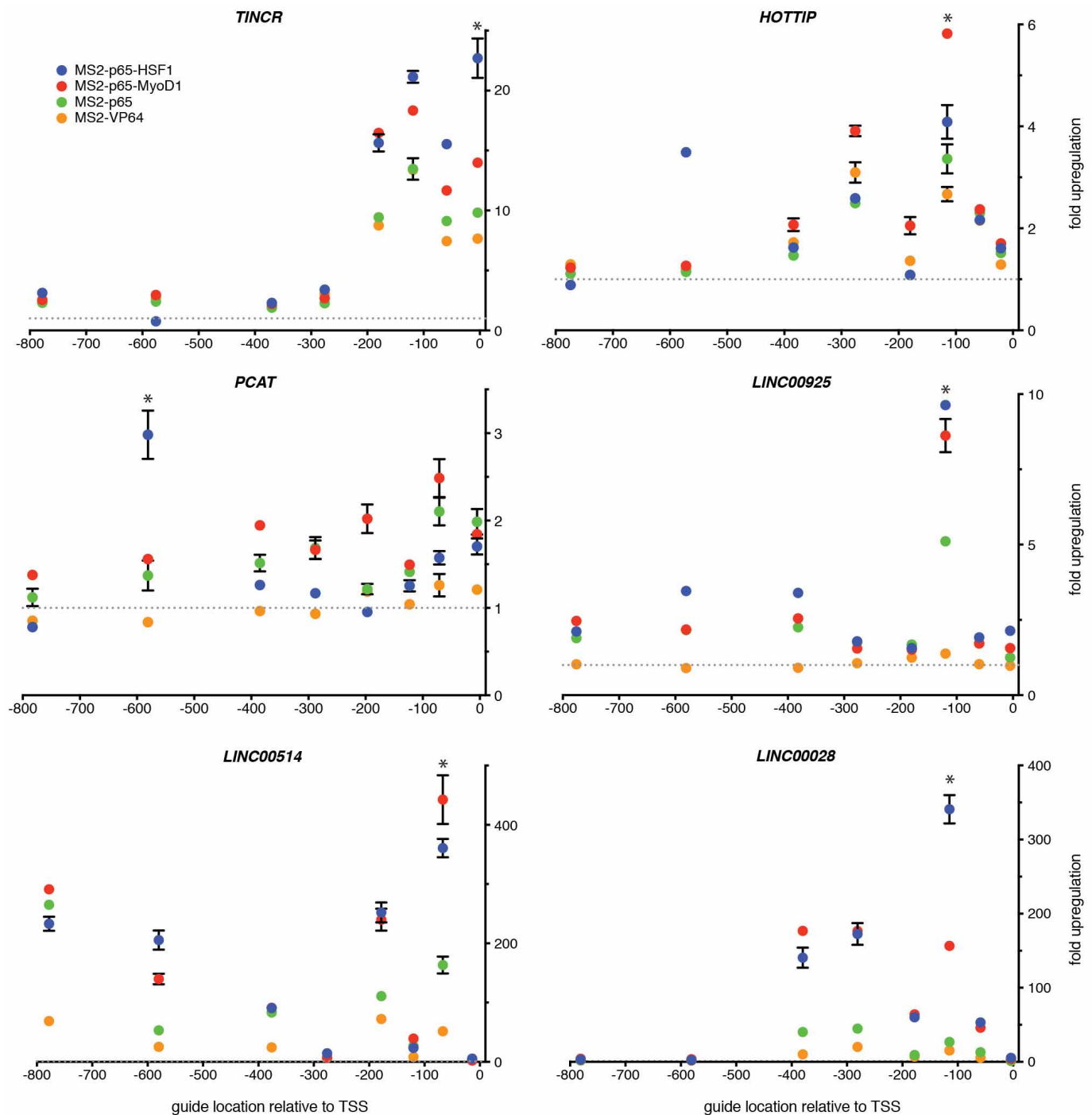
a, Schematic of the sgRNA stem loops showing contacts between each stem loop and Cas9. Contacting amino acid residues are highlighted in yellow. Tetraloop and stem loop 2 do not make any contacts with Cas9, whereas stem loops 1 and 3 share extensive contacts with Cas9. **b**, sgRNA 2.0 with MS2 stem loops inserted into the tetraloop and stem loop 2. **c**, Addition of a second NLS or an alternative HNH domain inactivating point mutation in Cas9 improve efficiency of transcription activation for *MYOD1* moderately. **d**, dCas9-VP64 activators exhibit improved performance by recruitment of

MS2-p65 to the tetraloop and stem loop 2. Addition of an AU flip or extension in the tetraloop does not increase the effectiveness of dCas9-mediated transcription activation. **e**, Tetraloop and stem loop 2 are amenable to replacement with MS2 stem loops. Base changes from the sgRNA 2.0 scaffold are shown at the respective positions, with dashes indicating unaltered bases and bases below dashes indicating insertions. Deletions are indicated by absence of dashes at respective positions. All figures are $n = 3$ and mean \pm s.e.m.



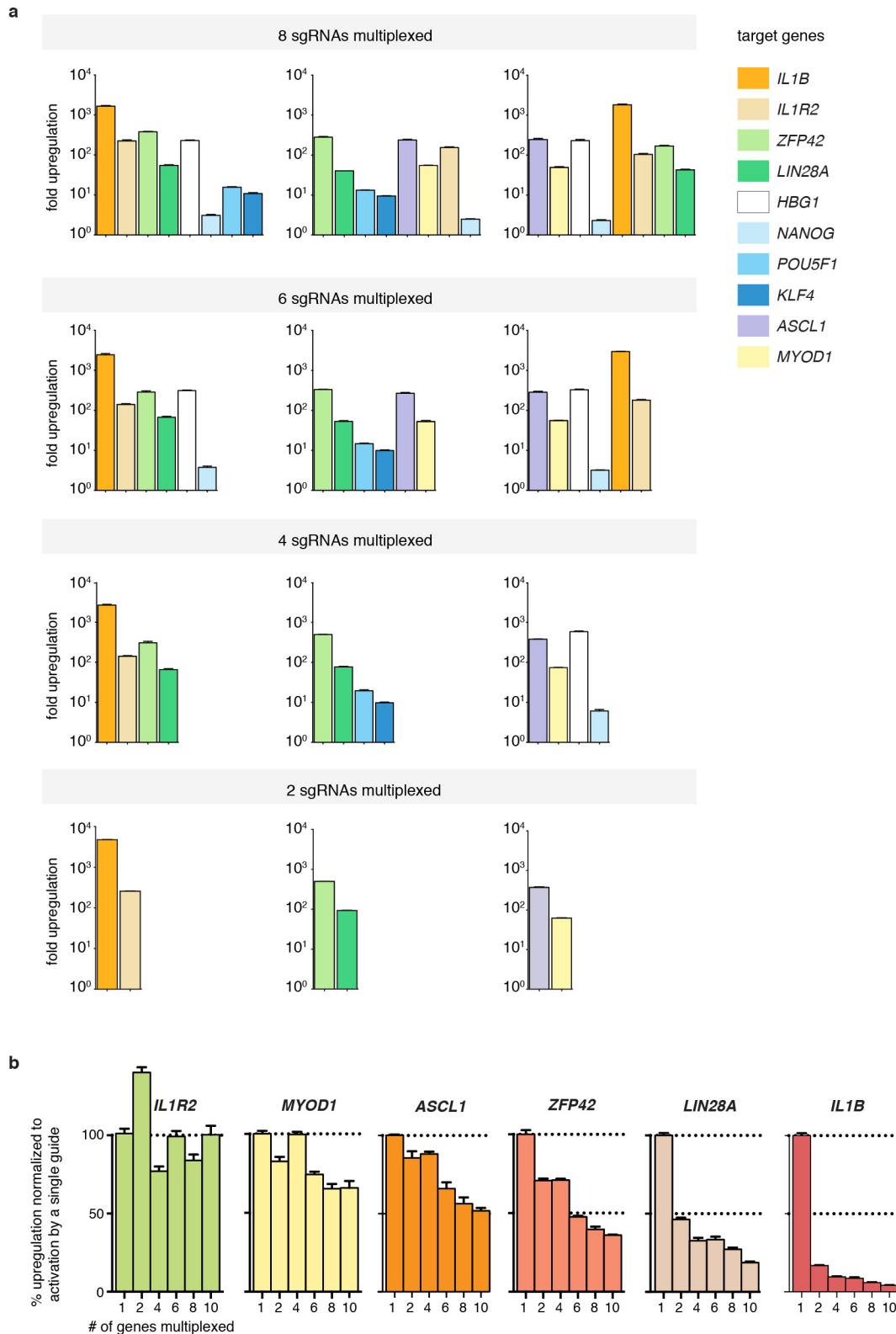
Extended Data Figure 2 | SAM mediates efficient activation of a panel of 12 coding genes and 6 lincRNAs. **a**, Comparison of the activation levels of 12 genes with dCas9-VP64 in combination with MS2-p65, MS2-p65-HSF1, or MS2-p65-MyoD1. MS2-p65-HSF1 mediated significantly higher levels of activation than MS2-p65 alone for 9 out of 12 genes. The best guide out of 8 tested for each gene (Fig. 2a) was used in this experiment. Activation levels for each type of MS2-fusion is presented as a percentage relative to the activation achieved using MS2-p65. **b**, Investigation of transcriptional changes in the

closest coding transcripts for SAM-mediated activation of 6 lincRNAs. Direction of the coding transcript relative to the lincRNA and distance between transcription start sites are shown. Only targeting of HOTTIP resulted in a significant change in the levels of the closest coding transcript (HOXA13). The best guide out of 8 tested for each gene (Fig. 2e) in combination with dCas9-VP64 and MS2-p65-HSF1 was used in this experiment. All figures are $n = 3$ and mean \pm s.e.m.



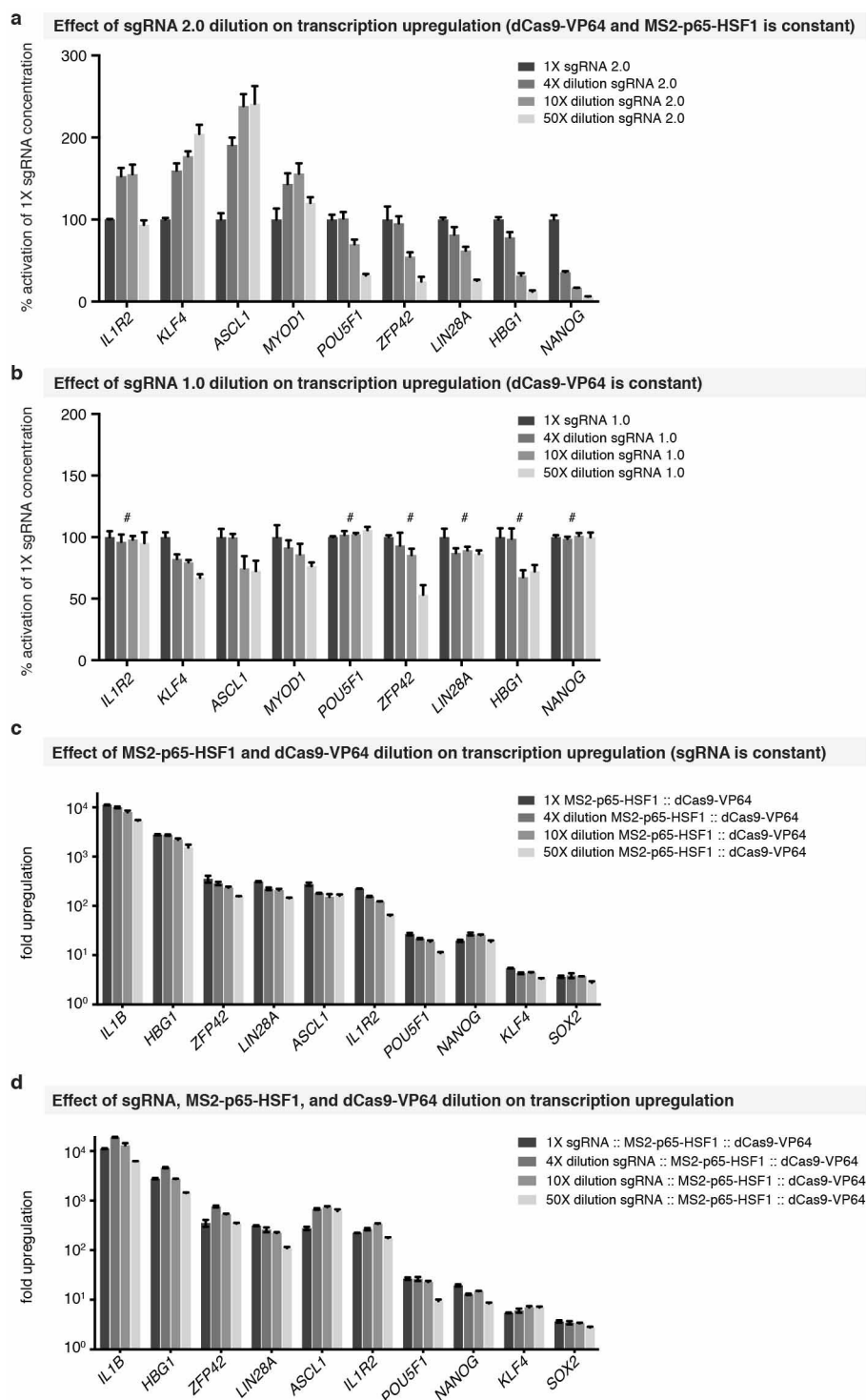
Extended Data Figure 3 | Activation of lincRNAs by SAM. Six lincRNAs, three characterized and three uncharacterized, were targeted using SAM. For each lincRNA, 8 sgRNAs were designed to target the proximal promoter region (+1 to -800 bp from the TSS) with 4 different MS2 activators (MS2-p65-HSF1, MS2-p65-MyoD1, MS2-p65, and MS2-VP64) in

combination with dCas9-VP64. MS2 activators with a combination of 2 different domains (MS2-p65-HSF1 or MS2-p65-MyoD1) consistently provided the highest activation for each lincRNA, $*P < 0.01$ for MS2-p65-HSF1 or MS2-p65-MyoD1 versus MS2-p65. $n = 3$ and mean \pm s.e.m. is shown.



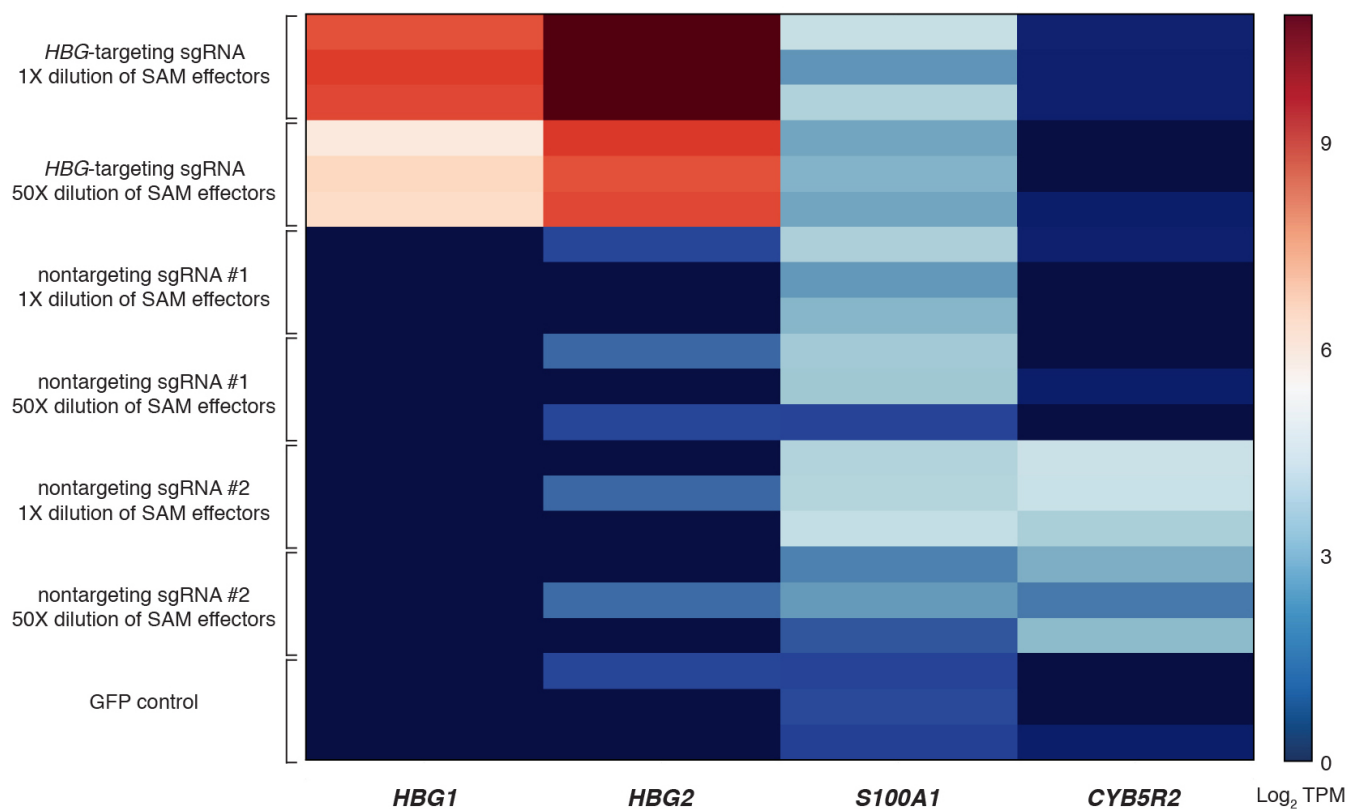
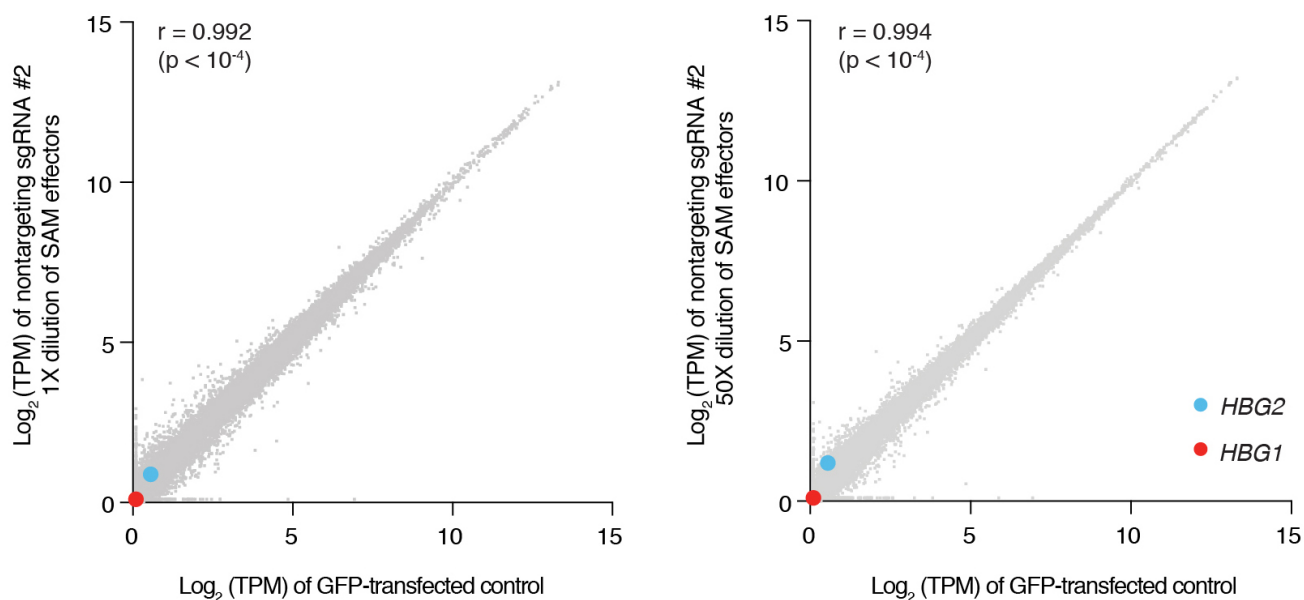
Extended Data Figure 4 | Multiplexed activation using SAM and activation of a panel of 10 genes as a function of SAM component dosage.
a, Activation of a panel of 10 genes by combinations of 2, 4, 6 or 8 sgRNAs simultaneously. The mean fold upregulation is shown on a log₁₀ scale.

MS2-p65-HSF1 and dCas9-VP64 were used in this experiment. **b**, The relative activation efficiency of individual sgRNAs varies depending on the target gene and the degree of multiplexing. $n = 3$ and mean \pm s.e.m. is shown.



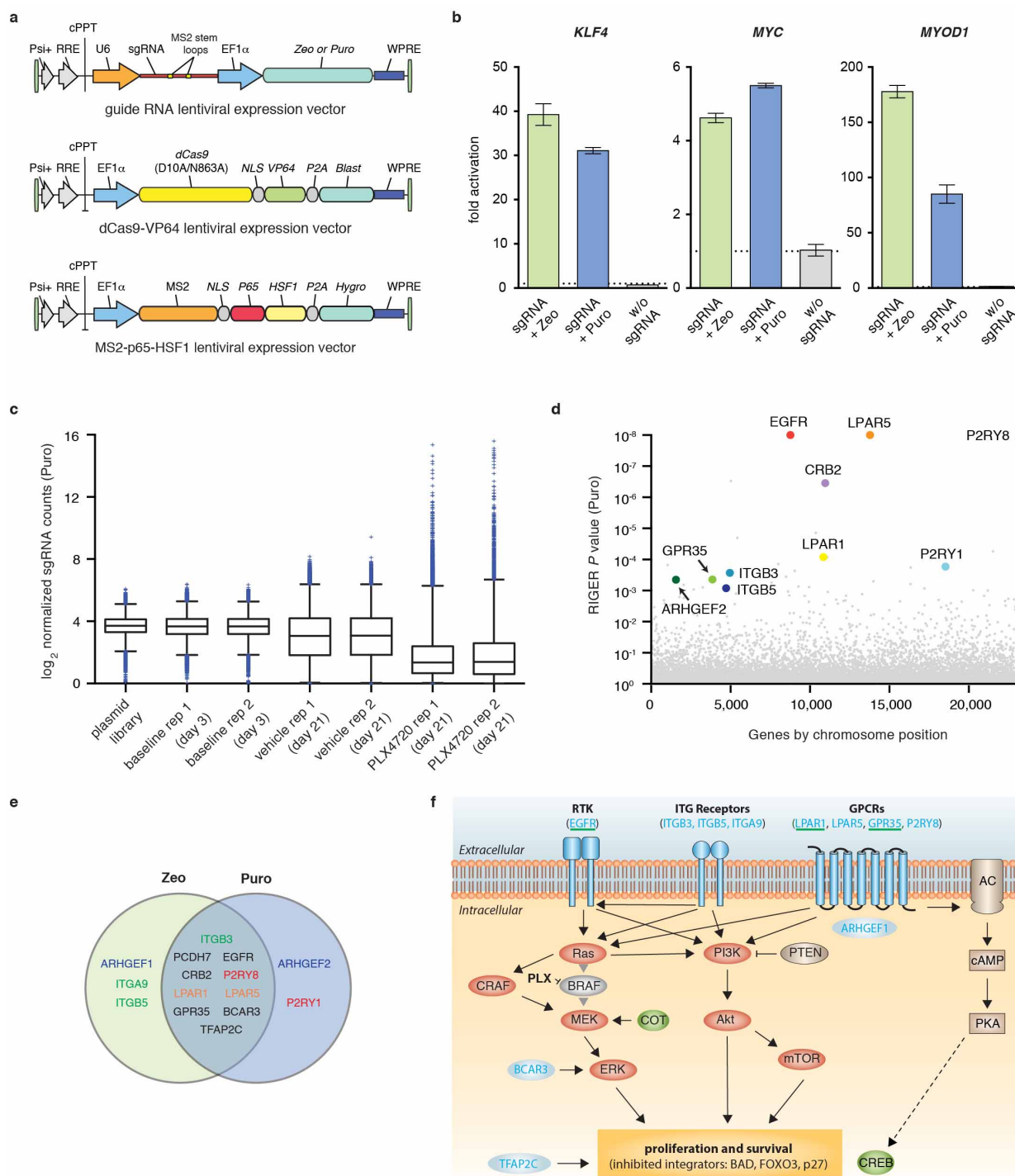
Extended Data Figure 5 | The effect of guide and SAM-component dilution on target activation. **a**, The results for dilution of sgRNA 2.0 on target activation. **b**, The result for dilution of sgRNA 1.0 on target activation. # denotes an activation of <twofold at 1× guide dilution. **c**, Effect of MS2-p65-HSF1 and dCas9-VP64 dilution, at 1:1, 1:4, 1:10 and 1:50 of the original dosage for each component, on the effectiveness of transcription upregulation. The

amount of sgRNA expression plasmid was kept constant. **d**, Effect of diluting all three SAM components (dCas9-VP64, MS2-p65-HSF1, and sgRNA) at 1:4, 1:10, and 1:50 of the original dosage for each component. Fold upregulation is calculated using GFP-transfected cells as the baseline. Error bars indicate s.e.m. and $n = 3$ for all figures.

a**b**

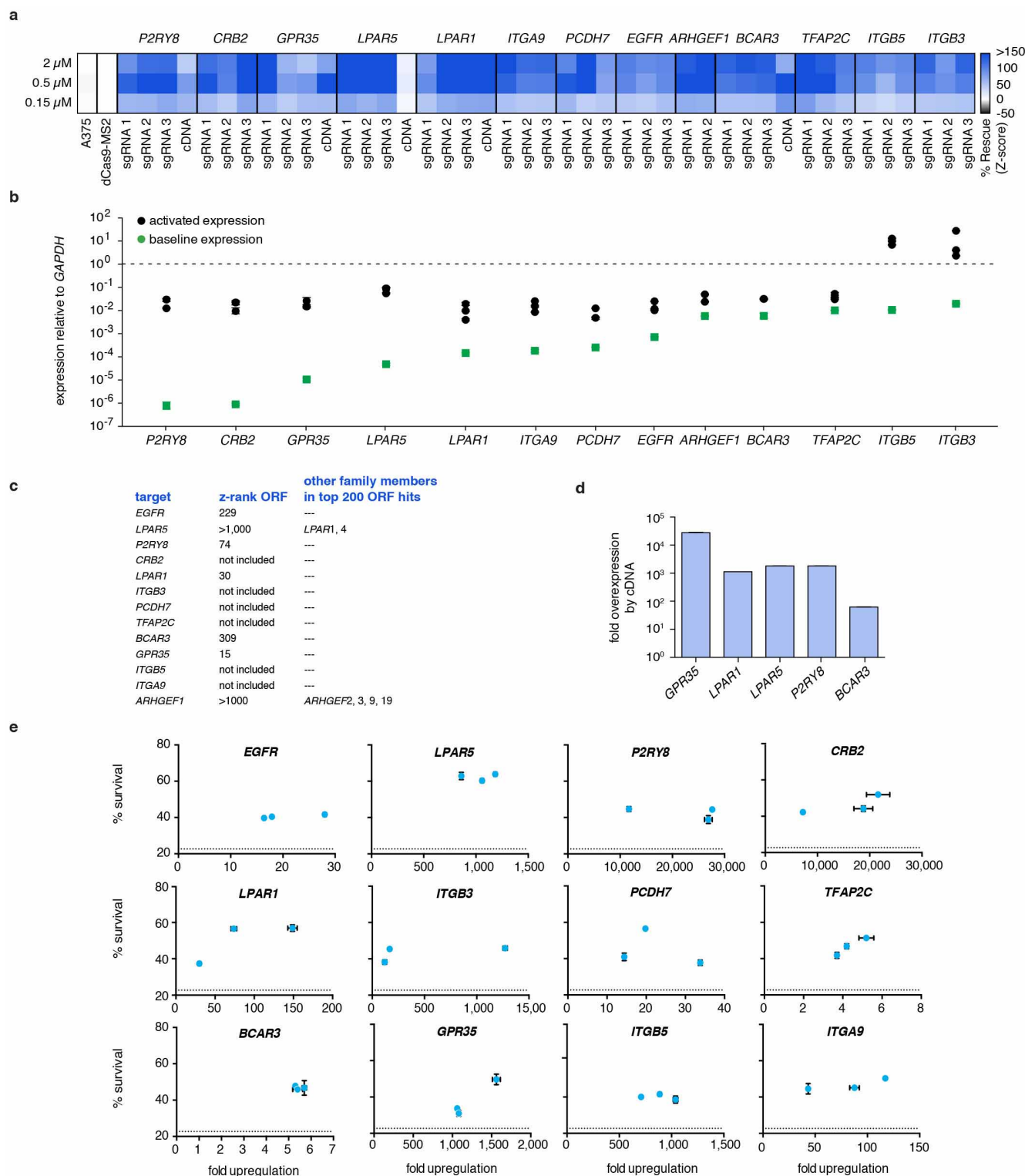
Extended Data Figure 6 | RNA-seq analysis of transcriptome changes mediated by SAM. **a**, A heat map of \log_2 (TPM) expression values of all statistically significant differentially expressed genes (t -test q value < 0.05 adjusted with FDR multiple hypothesis correction) found in any of the six experimental conditions compared to the GFP-transfected control.

b, Expression levels in \log_2 (TPM) values of all detected genes in RNA-seq libraries of GFP-transfected controls (x-axis of all graphs) compared to (from left to right): non-targeting control sgRNA no. 2 in 1X dilution and 50X dilution (y axis). Marked are *HBG1* (red) and *HBG2* (blue).



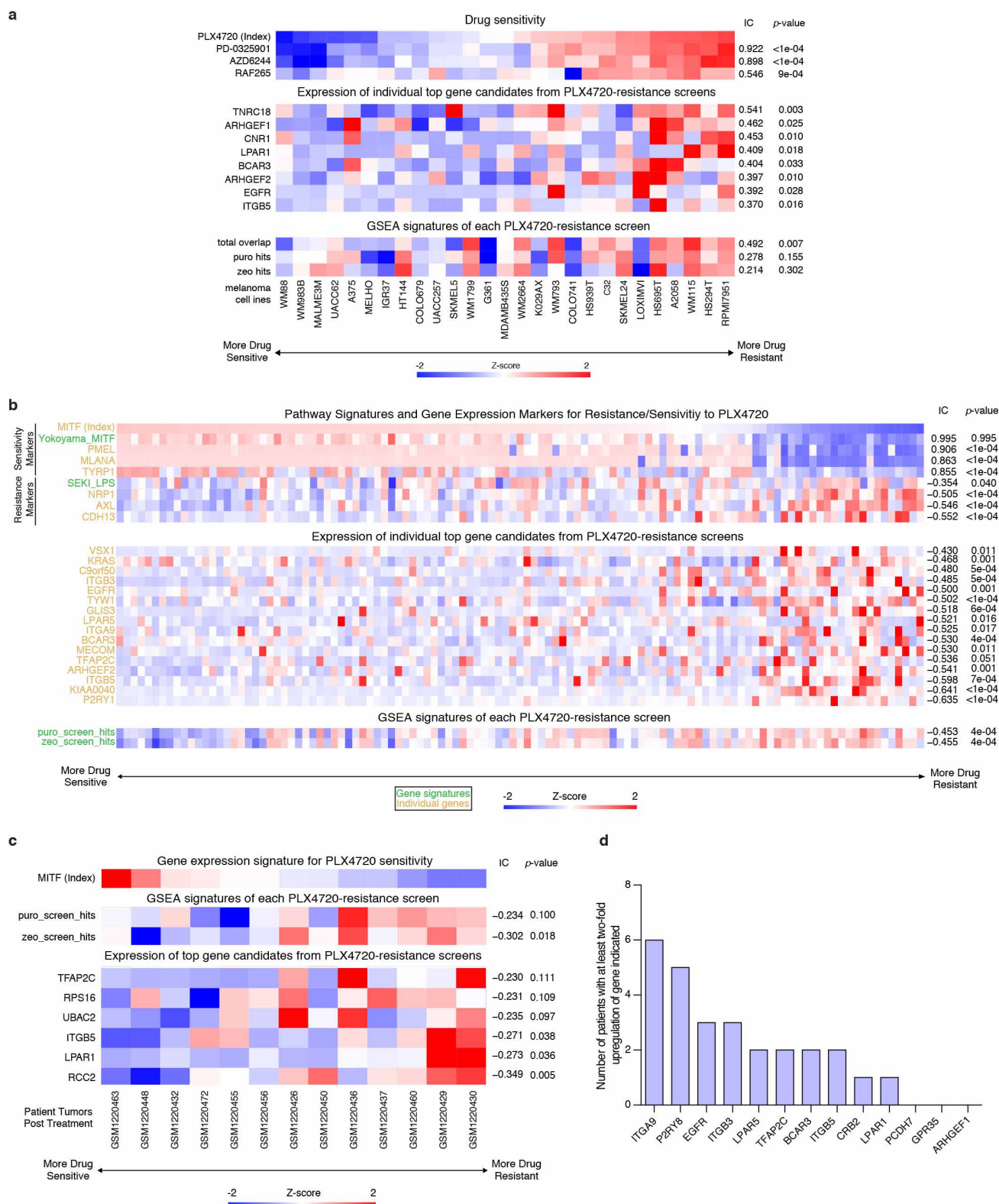
Extended Data Figure 7 | Genome-scale lentiviral screen using puromycin-resistant SAM sgRNA library. **a**, Design of three lentiviral vectors for expressing sgRNA, dCas9-VP64, and MS2-p65-HSF1. Each vector contains a distinct selection marker to enable co-selection of cells expressing all three vectors. **b**, Lentiviral delivery of SAM components was tested by first generating 293FT cell lines stably integrated with dCas9-VP64 and MS2-p65-HSF1, and subsequently transducing these cells with single-gene targeting lentiviral sgRNAs at MOI <0.2. Transcription activation efficiency is measured 4 days post sgRNA lentivirus transduction and selection with zeocin or puromycin. Activation is at least as effective as previously observed with transient transfection in all three cases. **c**, Box-plot showing the distribution of sgRNA frequencies at different time points post lentiviral transduction with the Puromycin library, after treatment with DMSO vehicle or PLX-4720. Two infection replicates are shown. **d**, Identification of top candidate genes using the RIGER *P* value analysis (KS method) based on the average of both infection

replicates. Genes are organized by positions within chromosomes. **e**, Overlap between the top 20 hits from the zeocin and puromycin screens. Genes belonging to the same family are indicated by the same colour. There is a 50% overlap between the top hits of each screen as shown in the intersection of the Venn diagram. **f**, Relevant signalling pathways in BRAF inhibitor resistance. Reactivation of the Ras-ERK pathway as well as the parallel PI3K-Akt pathway have previously been implicated as two alternative resistance mechanisms to BRAF inhibitors^{23,24,26-29}. Both pathways have been described as stimulating proliferation and survival⁴⁹. BAD, FOXO and p27 are common inhibited downstream targets⁴⁹. Recently, stimulation of the cAMP-CREB pathway by GPCRs has been described as a potential additional resistance mechanism³⁰. Top candidates from our screen are indicated in blue and putative connections to all three pathways are shown^{25,50,51}. Candidates previously validated to mediate PLX-4720 resistance are underlined in green^{26,30}. COT and CREB are independently validated mediators of resistance^{23,30}.



Extended Data Figure 8 | Individual validation of PLX-4720 resistance mediation by top screen hits. **a**, Validation of the top 10 Zeo screen hits and the top 10 shared hits (13 genes total). Every gene was independently activated by all three guides from the screen and tested for the ability to increase survival of A375 cells treated with three different concentrations of PLX-4720 (2 μ M, 0.5 μ M and 0.15 μ M). The z-score based on the % increase in survival relative to control (A375 cells transduced with dCas9-VP64 and MS2-p65-HSF1 alone) is shown for each guide and PLX-4720 concentration. Five cDNAs available from a previous large-scale gain-of-function PLX-4720 resistance screen were also included³⁰. Every guide for each top hit mediates significant PLX-4720 resistance. **b**, The same panel of top hits exhibits a large range of basal expression levels and is effectively activated by all guides. The

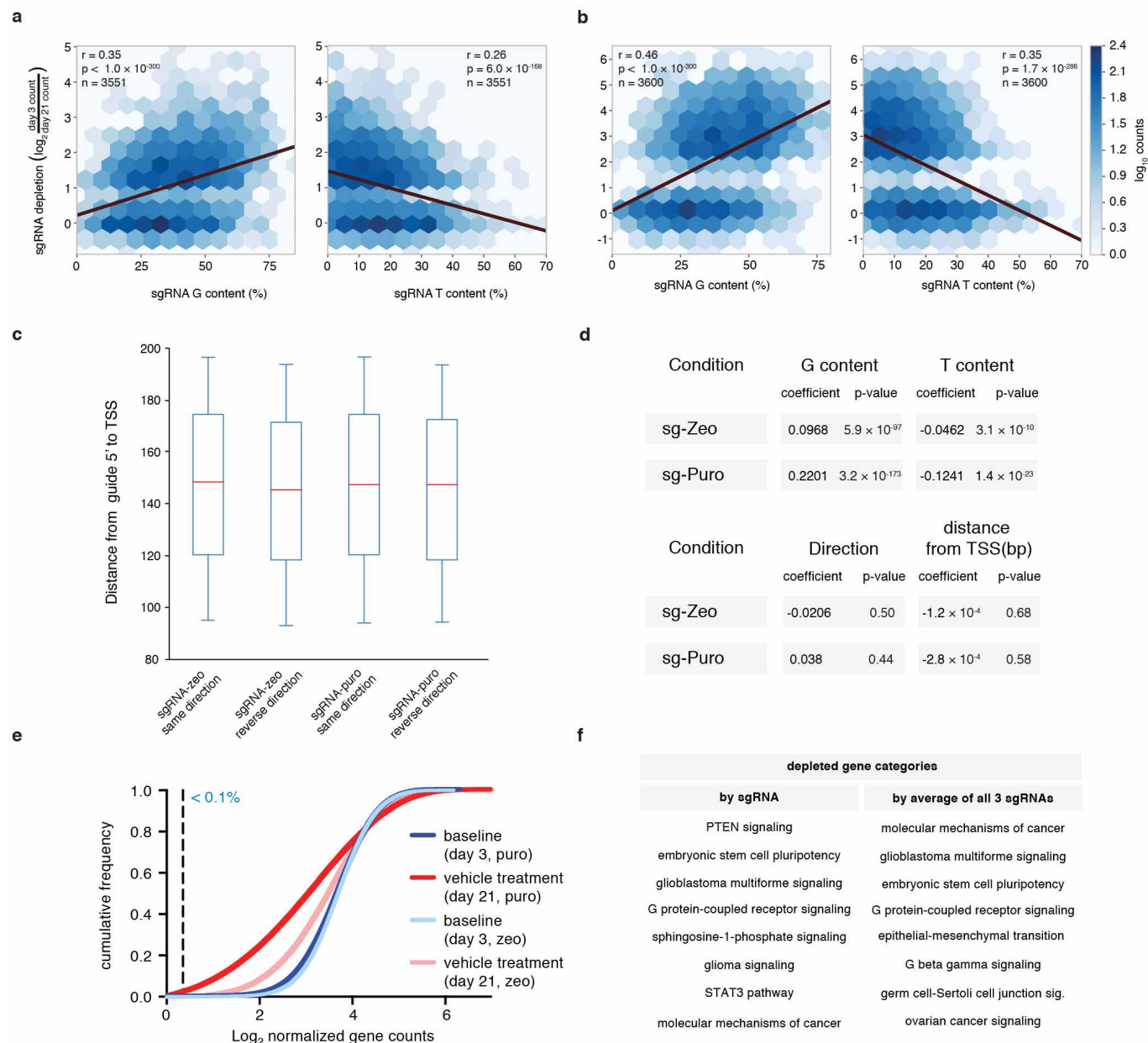
expression level relative to the housekeeping gene *GAPDH* is shown both at baseline as well as after activation by each individual guide. **c**, Ranks of the validated set of genes in the previous ORF screen. Six genes were not part of the cDNA library, five hits are shared (present in the top 3%) and only *LPAR5* and *ARHGEF1* were present but not highly ranked. Both of these genes had highly ranked members of the same family. **d**, Levels of overexpression from the five tested cDNA constructs. Transcript levels were higher for these five cDNAs than those mediated by SAM for the same genes. **e**, Correlation of survival at 2 μ M PLX-4720 treatment and transcript upregulation achieved by individual guides. For most genes (9 out of 12 shown), the percent survival is very similar across transcript levels achieved by all three guides. Dotted lines indicate control survival.



Extended Data Figure 9 | Expression of top hits and screen signatures are elevated in PLX-4720 resistant melanoma cell lines and patient samples.

a, Heat map showing sensitivity to different drugs (top), expression of SAM top screen hits (middle), and SAM screen signature scores (bottom; see Methods for signature generation) in Cancer Cell Line Encyclopedia cell lines³⁵. Drug sensitivities are measured as Activity Areas (AA). The melanoma cell lines are sorted by PLX-4720 drug sensitivity. RAF inhibitors: PLX-4720 and RAF265; MEK inhibitors: AZD6244 and PD-0325901. **b**, Heat map showing expression of gene/signature markers for BRAF-inhibitor sensitivity (top), expression of SAM top screen hits (middle) and screen signature scores

(bottom) in different BRAF(V600) patient melanoma samples (primary or metastatic) from The Cancer Genome Atlas. **c**, Heat map showing MITF expression (top), screen signature scores (middle), and expression of SAM top screen hits (bottom) in different BRAF(V600E) patient melanoma biopsies post-treatment with BRAF inhibitors³⁸. **d**, Bar chart showing the number of patients (out of 13 total) from **c** with at least a twofold change (post/pre-treatment) in gene expression of the top PLX-4720 screen hits in the post-treatment samples. All associations are measured using the information coefficient (IC) between the index and each of the features and *P* values are determined using a permutation test. All heat maps show Z scores.



Extended Data Figure 10 | Guide depletion analysis to identify gene set enrichment and guide efficiency parameters. **a, b,** Heat maps of sgRNA nucleotide content versus depletion after 21 days. sgRNA targeting significantly depleted genes (from RIGER analysis) in sgRNA-zeo (**a**) or sgRNA-puro (**b**) screens were analysed for trends based on G or T content in the sgRNA sequence. sgRNA depletion is positively correlated with G content and negatively correlated with T content. Other bases analysed (A and C) had significant ($P < 0.0007$) but weak ($r < 0.2$) negative correlation. **c,** 90% of guides analysed fall within a 100-bp window < 200 bp from the TSS. Boxplots of distance from 5' end of the guide to the TSS for sgRNA-zeo and sgRNA-puro in same and reverse direction (relative to target transcription). Whiskers span 5th to 95th quartile. **d,** Coefficients and P values for ordinary least squares predicting sgRNA depletion of significantly depleted genes from G content, T content, distance from 5' end of the guide to the TSS and direction of guide.

Only nucleotide content has a significant effect on depletion in this model, consistent with a high efficiency of guides within 200 bp of the TSS regardless of strand orientation (Fig. 2d). **e,** The cumulative frequency of sgRNAs 3 and 21 days after transduction in A375 cells is shown. Shift in the 21-day curve represents the depletion in a subset of sgRNAs. Less than 0.1% of all guides are not detected at day 3 (detected by less than 10 reads). **f,** Depleted guides (Supplementary Table 3) can be analysed for significant clustering of gene categories. Gene categories exhibiting significant depletion based on Ingenuity Pathway Analysis ($P < 0.01$ after Benjamini-Hochberg FDR correction) are shown. Categories based on the 1,000 most depleted guides individually (left) and the average of all 3 guides/gene (right). These categories include either positive or negative regulators of each pathway that reduce proliferation and survival.

A spin-down clock for cool stars from observations of a 2.5-billion-year-old cluster

Søren Meibom¹, Sydney A. Barnes^{2,3}, Imants Platais⁴, Ronald L. Gilliland⁵, David W. Latham¹ & Robert D. Mathieu⁶

The ages of the most common stars—low-mass (cool) stars like the Sun, and smaller—are difficult to derive^{1,2} because traditional dating methods use stellar properties that either change little as the stars age^{3,4} or are hard to measure^{5–8}. The rotation rates of all cool stars decrease substantially with time as the stars steadily lose their angular momenta. If properly calibrated, rotation therefore can act as a reliable determinant of their ages based on the method of gyrochronology^{2,9–11}. To calibrate gyrochronology, the relationship between rotation period and age must be determined for cool stars of different masses, which is best accomplished with rotation period measurements for stars in clusters with well-known ages. Hitherto, such measurements have been possible only in clusters with ages of less than about one billion years^{12–16}, and gyrochronology ages for older stars have been inferred from model predictions^{2,7,11,17}. Here we report rotation period measurements for 30 cool stars in the 2.5-billion-year-old cluster NGC 6819. The periods reveal a well-defined relationship between rotation period and stellar mass at the cluster age, suggesting that ages with a precision of order 10 per cent can be derived for large numbers of cool Galactic field stars.

Prior observations in star clusters with ages $\lesssim 300$ million years (Myr) have shown that cool stars begin their main-sequence phase with a dispersion in their rotation periods, P , spanning two orders of magnitude^{12–14,18} (0.1–10 d). However, this dispersion diminishes rapidly with cluster age, t , as they lose angular momentum through magnetically channelled winds¹⁹, causing their periods to increase and converge to a well-defined relationship with stellar mass, M , by the age (600 Myr) of the Hyades cluster^{15,20}. These observations suggest that cool main-sequence stars older than the Hyades probably occupy a single surface, $P = P(t, M)$, in P – t – M space, which can be defined by measurements of their periods and photometric colours (a proxy for stellar mass) in a series of age-ranked clusters (Fig. 1). Measuring stellar rotation periods in clusters older than the Hyades will confirm or deny the existence of such a surface, and, if it exists, define its shape and thickness.

Models of cool-star rotational evolution also describe a convergence of rotation with age^{2,11,17,21}. However, they differ in their predictions of the location and shape of the P – t – M surface and, beyond the common assumption that the rotation period of the Sun is typical for stars of its mass and age, were until 2011¹⁶ unconstrained at ages greater than 600 Myr. Observations to define the P – t – M surface at older ages are therefore required to extend our knowledge of the P – t – M relations and thereby allow the ages of individual cool stars in the field to be derived from their measured periods and colours by the method of gyrochronology¹¹.

The rotation period of a cool star can be determined from small ($\lesssim 1\%$) periodic modulations in its brightness as rotation carries star-spots across the stellar disc. Older stars have fewer and smaller spots, making their periods harder to detect. Accordingly, observations from ground-based telescopes have been unable to detect rotation periods in clusters older than the Hyades. Although periods are increasingly being

measured for isolated field stars²², their ages, unlike those of cluster stars, are not known to a precision adequate to calibrate gyrochronology.

With an age of 2.5 billion years (Gyr) (ref. 23), NGC 6819 bridges the large gap in age between the Sun and existing cluster observations (Fig. 1). The cluster was within the field of view of NASA's Kepler satellite, permitting a time-series photometric survey of its members by The Kepler Cluster Study¹⁶. Cool cluster members were selected for observation using prior ground-based photometry for the cluster²⁴, a 90 yr-baseline proper-motion study²⁵ and multi-epoch radial-velocity measurements over 15 yr (ref. 26 and an ongoing survey by S.M.) (Methods). The precision, cadence and duration of the Kepler photometry enable us to measure rotation periods for cool stars much older than the Hyades. Previous results from The Kepler Cluster Study in the 1 Gyr-old cluster NGC 6811 confirmed the existence of a unique relationship between P and M for cool stars at that age¹⁶, and measured a median rotation period of 10.8 d for solar-mass stars. Since then, a study to measure periods from Kepler data in NGC 6819 was carried out by another group²⁷. That work was limited to cluster stars of greater than solar mass, and was thus unable to define a P – M relationship for cool stars.

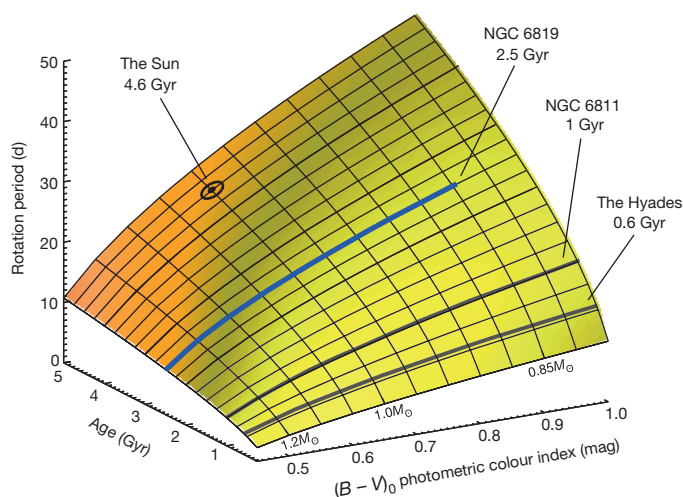


Figure 1 | The schematic P – t – M surface for cool stars. The hypothetical relationship between rotation period, age and colour extrapolated (yellow) to greater ages from the colour–period relations in young clusters using a particular P – t relationship⁵, and assuming that the Sun (marked by the black solar symbol; \odot) resides on it. The blue line indicates the locus of stars in NGC 6819 for which we have determined rotation periods. The dark grey lines at ages of 0.6 and 1 Gyr represent prior observations in the Hyades¹⁵ and NGC 6811¹⁶ clusters, respectively. Stellar masses in solar units are marked on the surface at the corresponding colours. (Figure adapted from ref. 16.)

¹Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA. ²Leibniz Institute for Astrophysics, An der Sternwarte 16, 14482 Potsdam, Germany. ³Space Science Institute, 4750 Walnut Street #205, Boulder, Colorado 80301, USA. ⁴Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA. ⁵Center for Exoplanets and Habitable Worlds, The Pennsylvania State University, University Park, Pennsylvania 16802, USA. ⁶Department of Astronomy, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA.

We have measured rotation periods for 30 cool stars in NGC 6819 (Extended Data Figs 1–5). All 30 stars are both photometric (from their location in the cluster's colour–magnitude diagram) and kinematic (from measurements of their proper motions and radial velocities) members of the cluster (Methods). Their de-reddened colour indices, $(B - V)_0$, range from 0.41 to 0.89 mag, corresponding to a stellar mass range from ~ 1.4 to 0.85 solar masses. Their periods range from 4.4 to 23.3 d and are displayed in the resulting colour–period diagram (CPD) for NGC 6819 (Fig. 2). Extended Data Table 1 lists all relevant properties for the 30 stars. The stars form a single and narrow sequence in the CPD. This sequence defines a clear dependence of increasing stellar rotation period (decreasing rotation rate) on increasing stellar colour (decreasing mass) and represents the cross-section of the hypothesized P – t – M surface at $t = 2.5$ Gyr (Fig. 1, blue line).

The solar-mass stars (defined here as those with $0.62 \text{ mag} \leq (B - V)_0 \leq 0.68 \text{ mag}$) all have periods between 17.36 and 18.70 d (mean, 18.2 d; s.d., 0.4 d), implying that the Sun's rotation period was probably in that range when it was the age of NGC 6819. With the 10.8 d median period for solar-mass stars in NGC 6811¹⁶, and the mean solar photometric rotation period of 26.1 d, this implies a Skumanich-type⁵ spin-down (P varies as $t^{1/2}$) for solar-mass stars over the 3.6 Gyr interval measured.

The relatively small number of periods detected for stars with $(B - V)_0$ from 0.47 to 0.57 mag does not reflect a lack of cluster members (Extended Data Fig. 6). A similar pattern was seen for NGC 6811¹⁶, and is unsurprising to photometrists, who know that such stars show little variability. The colour index $(B - V)_0 = 0.47 \text{ mag}$ separates stars with radiative envelopes from those with convective envelopes and is associated with the onset of effective magnetic wind braking (the 'break in the Kraft curve'²⁸). The rotation periods of the more massive stars in NGC 6819 ($(B - V)_0 < 0.47 \text{ mag}$) are scattered around a median of 4.8 d, demonstrating a steeper spin-down (P varies as t) from a median period of 1.3 d in the 1 Gyr-old cluster NGC 6811¹⁶.

The 30 rotation periods were determined by Lomb–Scargle periodogram analysis²⁹ of long-cadence (30 min exposures) Kepler light curves spanning ~ 3.75 yr (Methods). The rotation period for a given star was

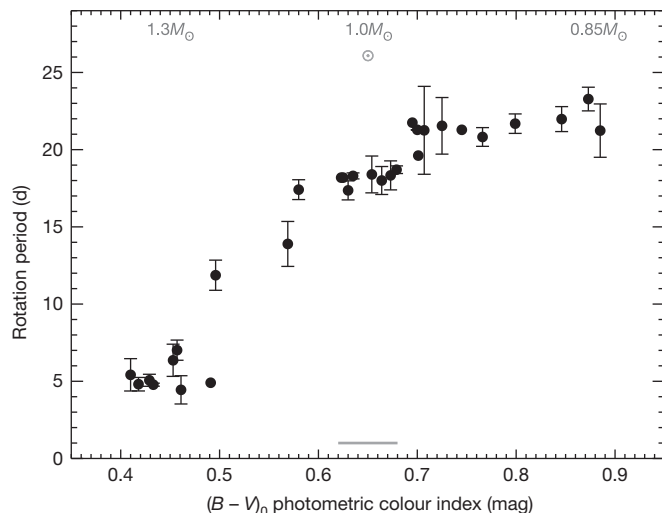


Figure 2 | The colour–period diagram for NGC 6819. The distribution of rotation periods as a function of de-reddened colour index $(B - V)_0$ for 30 cool photometric, proper-motion, and radial-velocity members of the 2.5 Gyr open star cluster NGC 6819. The measurements define a tight dependence of rotation period on colour (mass). The symbols and error bars respectively indicate the means and standard deviations of multiple measurements for the same star when available. The location of the Sun (4.56 Gyr) in the diagram is marked with a grey solar symbol. Stellar masses in solar units are given along the top horizontal axis at the corresponding colours. Solar-mass stars with $(B - V)_0$ between 0.62 and 0.68 mag (interval marked by grey line near the bottom horizontal axis) have a mean period of 18.2 d with a standard deviation of 0.4 d.

determined from subsections of the full light curve, chosen to avoid and minimize the effect on the measured period of multiple spots or spot groups, or that of trends not removed by the data processing, or both. For all periods reported we have manually examined the periodogram and the phased and unphased light curves, determined the periods independently using the CLEAN algorithm³⁰, and assessed the level of contamination from neighbouring stars (Methods and Extended Data Fig. 7).

We have also measured the projected rotation velocities ($v \sin(i)$, where v is the stellar rotation velocity and i is the inclination angle between the stellar spin axis and the observer's line of sight) spectroscopically for 25 of the 30 stars. These $v \sin(i)$ values are fully consistent with the photometric rotation periods (Extended Data Fig. 8). The resolution of the spectra for the five remaining stars is too low to provide meaningful constraints on their rotation velocities.

The measured rotation periods in NGC 6819 establish that the P – t – M surface is well defined at ages beyond 1 Gyr. Together with prior observations in younger clusters, they specify the location, shape and thickness of the surface to an age of 2.5 Gyr.

The CPD is the projection of this surface onto the colour–period plane. Therefore, in this diagram, the measured rotational sequence for NGC 6819 separates stars younger (below the sequence) and older (above the sequence) than 2.5 Gyr, independent of any theoretical model of stellar rotational evolution. Because angular momentum loss in cool single stars is driven by internal processes independent of environment (pathological tidally interacting systems are an exception to this, but tidal effects on rotation are not a concern for $\sim 97\%$ of cool field stars; see Methods), it is the same for both cluster and field stars. Therefore, this classification, and, more generally, the relationship between rotation and age, must also be valid for cool field stars.

The relative scatter ($\Delta P/P$) about the NGC 6819 rotational sequence is $\sim 10\%$ for our entire sample, $\sim 5\%$ for stars with $(B - V)_0 > 0.55 \text{ mag}$ and $\sim 2\%$ (0.4 d) for the domain surrounding solar-mass stars that we are able to define particularly well. This scatter includes contributions from period measurement uncertainties, and the residual effects of both stellar differential rotation and the spread in initial periods on the 'zero-age main sequence'. Its relatively small size demonstrates that these effects do not prevent the determination of age from rotation, and that the P – t – M surface is intrinsically thin at this age, implying that ages determined from spin-down ('gyro ages') will be precise.

The derivation of such ages requires a model. A number of such models exist^{2,7,11,17}, individually differing with respect to the functional forms of the underlying variables, and even with respect to what those specific variables are. The NGC 6819 rotation period data permit a comparison between the predictions of these rotational evolution models for its age and the actual measurements. Figure 3 shows this comparison.

The (unaltered) model of gyrochronology from ref. 11 provides a good fit to the data over nearly the full colour range of the observations. To test the precision of gyrochronology we may thus treat the coeval NGC 6819 stars as individual field stars, and ask what age the model of ref. 11 would provide for each of the 21 best-measured stars, that is, those with $(B - V)_0$ colour index between 0.55 and 0.9 mag (masses between ~ 1.1 and ~ 0.85 that of the Sun). Every one of these stars returns a gyro age between 2 and 3 Gyr, with a roughly Gaussian distribution centred at 2.49 Gyr (Extended Data Fig. 9). The standard deviation of the 21 ages is 0.25 Gyr (10% of the mean gyro age), implying that ages of this precision can be derived for similarly well-measured field stars, despite the effects of measurement errors, differential rotation and a spread in initial rotation periods.

The mean age of 2.49 Gyr also represents the gyro age for NGC 6819. The standard error in this cluster age is 0.056 Gyr (that is, a 2% uncertainty, ignoring possible systematic errors in gyrochronology). The cluster gyro age thus agrees to within the uncertainty with the classical stellar evolution age of the cluster²³, implying that gyrochronology is well-calibrated at 2.5 Gyr.

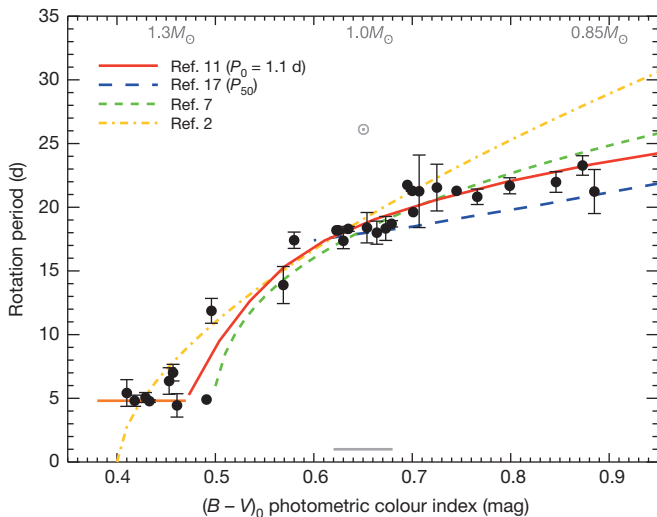


Figure 3 | Comparison between gyrochronology models and the NGC 6819 CPD. The predictions from four different models^{2,7,11,17} of cool star rotation periods at 2.5 Gyr are plotted against the measured periods in NGC 6819. All plotted models predict an observed increase in rotation period with increasing $(B - V)_0$ colour (decreasing stellar mass). The colour–period relation from ref. 11 fits the observations for stars with $(B - V)_0 > 0.55$ mag. The symbols and error bars respectively indicate the means and standard deviations of multiple measurements for the same star when available. The location of the Sun (4.56 Gyr) is marked with a grey solar symbol. Stellar masses in solar units are given along the top horizontal axis at the corresponding colours. The colour range for solar-mass stars is marked with a solid grey line near the bottom horizontal axis. The orange horizontal line for $(B - V)_0 < 0.47$ mag marks the median period of 4.8 d for stars in this colour range (~ 1.2 – 1.4 solar masses). P_0 in the model from ref. 11 refers to the initial (zero-age main sequence) rotation period. P_{50} in the model from ref. 17 refers to the 50th percentile rotation period for a given stellar mass.

We conclude that gyrochronology can provide accurate and precise ages for large numbers of cool stars with measured rotation periods. Such ages will enable us to study how astrophysical phenomena involving cool stars evolve over time, and will therefore be important to a wide range of research from the Galactic scale down to the scale of individual stars and their companions.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 October; accepted 27 November 2014.

Published online 5 January 2015.

- Soderblom, D. R. The ages of stars. *Annu. Rev. Astron. Astrophys.* **48**, 581–629 (2010).
- Barnes, S. A. Ages for illustrative field stars using gyrochronology: viability, limitations, and errors. *Astrophys. J.* **669**, 1167–1189 (2007).
- Sandage, A. The ages of M67, NGC 188, M3, M5, and M13 according to Hoyle's 1959 models. *Astrophys. J.* **135**, 349–365 (1962).
- Demarque, P. R. & Larson, R. B. The age of galactic cluster NGC 188. *Astrophys. J.* **140**, 544–551 (1964).
- Skumanich, A. Time scales for CA II emission decay, rotational braking, and lithium depletion. *Astrophys. J.* **171**, 565–567 (1972).
- Noyes, R. *et al.* Rotation, convection, and magnetic activity in lower main-sequence stars. *Astrophys. J.* **279**, 763–777 (1984).
- Mamajek, E. E. & Hillenbrand, L. A. Improved age estimation for solar-type dwarfs using activity-rotation diagnostics. *Astrophys. J.* **687**, 1264–1293 (2008).
- Jeffries, R. D. Using rotation, magnetic activity and lithium to estimate the ages of low mass stars. In *Proc. 23rd Evry Schatzman School on Stellar Astrophysics, 'The Ages of Stars'* (eds Charbonnel, C. *et al.*) 289–325 (EAS, 2014); preprint at <http://arxiv.org/abs/1404.7156v2> (2014).
- Kawaler, S. D. Rotational dating of middle-aged stars. *Astrophys. J.* **343**, L65–L68 (1989).
- Barnes, S. A. On the rotational evolution of solar- and late-type stars, its magnetic origins, and the possibility of stellar gyrochronology. *Astrophys. J.* **586**, 464–479 (2003).

- Barnes, S. A. A simple nonlinear model for the rotation of main-sequence cool stars. I. Introduction, implications for gyrochronology, and color-period diagrams. *Astrophys. J.* **722**, 222–234 (2010).
- Meibom, S. *et al.* Stellar rotation in M35: mass-period relations, spin-down rates, and gyrochronology. *Astrophys. J.* **695**, 679–694 (2009).
- Hartman, J. D. *et al.* A large sample of photometric rotation periods for FGK Pleiades stars. *Mon. Not. R. Astron. Soc.* **408**, 475–489 (2010).
- Meibom, S. *et al.* The color-period diagram and stellar rotational evolution—new rotation period measurements in the open cluster M34. *Astrophys. J.* **733**, 115–129 (2011).
- Delorme, P. *et al.* Stellar rotation in the Hyades and Praesepe: gyrochronology and braking time-scale. *Mon. Not. R. Astron. Soc.* **413**, 2218–2234 (2011).
- Meibom, S. *et al.* The Kepler Cluster Study: stellar rotation in NGC 6811. *Astrophys. J.* **733**, L9–L14 (2011).
- Epstein, C. R. & Pinsonneault, M. H. How good a clock is rotation? The stellar rotation-mass-age relationship for old field stars. *Astrophys. J.* **780**, 159–183 (2014).
- van Leeuwen, F. *et al.* VBLUW observations of Pleiades G and K dwarfs. *Astron. Astrophys.* **67** (suppl.), 483–506 (1987).
- Parker, E. N. Dynamics of the interplanetary gas and magnetic fields. *Astrophys. J.* **128**, 664–676 (1958).
- Radick, R. R. *et al.* The activity, variability, and rotation of lower main-sequence Hyades stars. *Astrophys. J.* **321**, 459–472 (1987).
- Kawaler, S. D. Angular momentum loss in low-mass stars. *Astrophys. J.* **333**, 236–247 (1988).
- McQuillan, A. *et al.* Rotation periods of 34,030 Kepler main-sequence stars: the full autocorrelation sample. *Astrophys. J.* **711**, 24–38 (2014).
- Jeffries, M. W. *et al.* WOCs 40007: a detached eclipsing binary near the turnoff of the open cluster NGC 6819. *Astron. J.* **146**, 58–78 (2013).
- Kalirai, J. S. *et al.* The CFHT Open Star Cluster Survey. II. Deep CCD photometry of the old open star cluster NGC 6819. *Astron. J.* **122**, 266–282 (2001).
- Platais, I. *et al.* WIYN Open Cluster Study. LV. Astrometry and membership in NGC 6819. *Astron. J.* **146**, 43–56 (2013).
- Milliman, K. E. *et al.* WIYN Open Cluster Study. LX. Spectroscopic binary orbits in NGC 6819. *Astron. J.* **148**, 38–58 (2014).
- Balona, L. A. *et al.* Kepler observations of the open cluster NGC 6819. *Mon. Not. R. Astron. Soc.* **430**, 3472–3482 (2013).
- Kraft, R. P. Studies of stellar rotation. V. The dependence of rotation on age among solar-type stars. *Astrophys. J.* **150**, 551–569 (1967).
- Scargle, J. D. Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. *Astrophys. J.* **263**, 835–853 (1982).
- Roberts, D. H. *et al.* Time series analysis with CLEAN - Part one - Derivation of a spectrum. *Astron. J.* **93**, 968–989 (1987).

Acknowledgements S.M. acknowledges support through NASA grant NNX09AH18A (The Kepler Cluster Study), NSF grant 1312882 (The Kepler Cluster Study: Planets and Gyrochronology) and the Smithsonian Institution's Competitive Grants Program for Science in 2012 and 2013. S.A.B. acknowledges support from the German Science Foundation (DFG) during a crucial phase of this work via a Mercator Guest Professorship at the University of Potsdam and the Leibniz Institute for Astrophysics Potsdam, Germany. This paper includes data collected by the Kepler mission. Kepler was competitively selected as the tenth Discovery mission. Funding for the Kepler mission is provided by the NASA Science Mission directorate. Some or all of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NNX13AC07G and by other grants and contracts. Spectroscopic observations of NGC 6819 with the Hectochelle spectrograph were obtained at the MMT Observatory, a joint facility of the Smithsonian Institution and the University of Arizona.

Author Contributions S.M. is the Principal Investigator for The Kepler Cluster Study and led the planning and execution of the study; the determination and validation of rotation periods from the Kepler data; the membership, binarity and vsin(*i*) survey in NGC 6819 with the Hectochelle spectrograph on the MMT telescope; and the writing of this paper. S.A.B. is a Co-Investigator on The Kepler Cluster Study and participated in planning the study, evaluated the light curves for periodicity, performed the gyrochronology analysis and age determination for the stars, and collaborated closely with S.M. in writing the paper. I.P. is a Co-Investigator on The Kepler Cluster Study and contributed to the selection of Kepler targets in NGC 6819 with proper-motion membership information and to the analysis of crowding and contamination in the vicinity of target stars from deep, high-spatial-resolution images of this star cluster. R.L.G. is a Co-Investigator on The Kepler Cluster Study and contributed to the selection of Kepler targets in NGC 6819 with analysis of crowding and contamination in deep, high-spatial-resolution images. He carried out analysis of alternative methods for extracting light curves from raw Kepler data. D.W.L. participated in the NGC 6819 radial-velocity membership and binarity surveys, and led the preparation of the Kepler Input Catalog. R.D.M. is the Principal Investigator of the WIYN Open Cluster Study, which contributed radial-velocity measurements to the membership and binary star surveys in NGC 6819.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.M. (smeibom@cfa.harvard.edu).

METHODS

1. Cluster membership for the 30 stars with measured rotation periods

1.1. Radial-velocity membership. The common space motion of the stars in a cluster is an effective way to distinguish them from foreground or background stars in the Galactic disk. Using the Hydra and Hectochelle multi-object spectrographs on the WIYN 3.5 m and MMT 6.5 m telescopes we have measured radial (line-of-sight) velocities over 15 yr for more than 4,300 stars within a circular 1°-diameter field centred on NGC 6819 (ref. 26 and an ongoing survey by S.M.). The Hydra spectra cover a 25 nm window centred on 513 nm and have a resolution of $\sim 20,000$. The Hectochelle spectra have resolutions of $\sim 40,000$ over a 15 nm window centred at 522 nm. For the late-type stars in NGC 6819 these spectral ranges are rich with narrow absorption lines and are thus well suited for radial-velocity measurements. Our radial-velocity measurement precision for stars of spectral types F, G and K (masses from ~ 1.3 to 0.7 solar masses) is $\sim 0.4 \text{ km s}^{-1}$ for stars brighter than 18.5 mag in the V band.

Against the broad radial-velocity distribution of Galactic field stars in the direction of the cluster, the members of NGC 6819 populate a distinct peak with a mean radial velocity of $+2.6 \pm 0.8 \text{ km s}^{-1}$. The uncertainty represents the velocity dispersion among the stars caused by internal cluster dynamics, binary orbital motions and observational errors. For a given star, the probability of cluster membership (P_{RV}) is calculated from simultaneous fits of separate Gaussian functions to the cluster (FC) and field (FF) radial-velocity distributions. The probability is defined as the ratio of the cluster-fitted value over the sum of the cluster- and field-fitted values at the star's radial velocity²⁶ (RV):

$$P_{\text{RV}} = \text{FC}(\text{RV}) / [\text{FC}(\text{RV}) + \text{FF}(\text{RV})]$$

The 30 stars in NGC 6819 with measured rotation periods are all radial-velocity members of the cluster ($P_{\text{RV}} > 50\%$) and their membership probabilities are given in Extended Data Table 1. The radial-velocity measurements also suggest that none of the 30 stars are in short-period binary stars, and, thus, that their angular momentum evolution is not affected by tidal interactions^{31,32}.

1.2. Proper-motion membership. For NGC 6819, a number of archival photographic plates are available, taken between 1919 and 1973 with long-focus telescopes. A total of 23 photographic plates were digitized using the Space Telescope Science Institute's GAMMA II multi-channel microdensitometer³³. Combining the measurements of these plates with matching second-epoch high-spatial-resolution CCD images, obtained in 2009 with the MegaCam camera on the 3.6 m Canada-France Hawaii Telescope (CFHT), allows us to derive accurate proper motions for stars in the field of NGC 6819. Proper motions were calculated for 15,750 stars down to 22 mag in V over a $40 \text{ arcmin} \times 40 \text{ arcmin}$ field centred on NGC 6819. The accuracy of the proper motions for well-measured stars is $\sim 0.2 \text{ mas yr}^{-1}$. For the inner parts (within a 15 arcmin radius from the centre) of the cluster this accuracy holds for stars brighter than 18 mag in V. Considering that the intrinsic dispersion of field star proper motions in the direction of NGC 6819 is about 3 mas yr^{-1} , this high accuracy enables a clean separation of cluster stars from field stars. Cluster membership probabilities (P_{μ}) were calculated using

$$P_{\mu} = \Phi_{\text{Cluster}} / (\Phi_{\text{Cluster}} + \Phi_{\text{Field}})$$

where Φ_{Cluster} and Φ_{Field} are the two-dimensional Gaussian frequency distributions of the cluster and field stars, respectively²⁵.

The majority (21) of the 30 stars with measured rotation periods have $P_{\mu} > 90\%$. For the remaining stars P_{μ} is lower because they are either located in the periphery of the cluster or their proper-motion errors are higher than expected at their apparent magnitudes²⁵. The latter usually is due to some degree of stellar image overlap.

1.3 Photometric membership. The colour-magnitude diagram (CMD; Extended Data Fig. 6) provides a third set of criteria for cluster membership. In the CMD, cluster members trace a well-defined relationship between stellar mass ($B - V$ colour index) and luminosity (brightness, V). Extended Data Fig. 6 shows V and $(B - V)_0$ for proper-motion members of NGC 6819. The cluster members form a clearly visible diagonal band in the CMD and a 'hockey-stick'-like turn-off near $(B - V)_0 \approx 0.45 \text{ mag}$ and $V \approx 16 \text{ mag}$. The locations of the 30 stars with measured rotation periods are marked with larger red circles. They are located on the cluster band in the CMD, making them photometric members of NGC 6819. The two members near $(B - V)_0 = 0.49 \text{ mag}$ are photometric binaries. The combined light from the two stars in the binary system places them above the cluster sequence. For these two stars, three and four radial-velocity measurements over 1,732 d and 1,074 d, respectively, show radial-velocity variations near our measurement precision, suggesting that these stars are members of relatively wide binary systems.

2. Data and data analysis. Stars identified as members of NGC 6819 were added to the list of targets for the Kepler mission as part of The Kepler Cluster Study¹⁶. The 30 stars for which we measure rotation periods were observed by Kepler for

2.5 yr, on average, over $\sim 3.75 \text{ yr}$. NGC 6819 was located in the part of the Kepler field of view covered by the CCD module that failed in January 2010. Since then, targets in NGC 6819 could be observed for only three of the four quarters each year.

Stellar rotation periods were derived from Kepler data summed into long-cadence ($\sim 30 \text{ min}$) bins. The data were processed by version 8.0 of the Kepler mission's data analysis pipeline and corrected by the Kepler Presearch Data Conditioning (PDC) module of the pipeline with an additional Bayesian maximum a posteriori (MAP) approach for the removal of systematics while preserving astrophysical signals such as rotational modulation^{34,35}. Before performing our period search, all quarters of the corrected data were normalized by the median signal and joined together to form a single light curve for each star. We used Lomb-Scargle periodogram analysis²⁹ to detect periodic variability, searching 20,000 frequencies corresponding to periods between 0.05 and 100 d. The rotation period for a given star was determined from between one and seven separate time intervals distributed over the full light curve. The median peak-to-peak amplitude of variability for all period detections is 4 mmag, with a range of 1 to 125 mmag. When more than one rotation period measurement was possible for a given star, the mean period was calculated and used. For all stars, the periodogram and the raw and phased light curves were examined by eye. The periods were derived independently by both S.M. and S.A.B. using different analysis tools and algorithms. Extended Data Figs 1–5 shows examples, for all 30 stars, of PDC-MAP corrected light curve intervals used to measure their rotation periods. The figure also shows the light curves phase-folded on the periods, and the corresponding periodogram for each star. Extended Data Table 1 lists, for each of the 30 stars, basic astrometric and photometric data, the radial-velocity and proper-motion cluster membership probabilities, the number of period measurements, and their period mean and standard deviation.

For 25 of the 30 stars, spectra acquired with Hectochelle enabled determination of their projected rotation velocities ($v \sin(i)$) via cross-correlation with a library of synthetic spectra. Extended Data Fig. 8 shows the mean rotation periods versus the mean $v \sin(i)$ for stars in NGC 6819. It also displays three curves tracing the expected relation between rotation period and rotation velocity for a 90° inclination angle (i) of the stellar rotational axis and stellar radii of 0.85, 1.0, 1.4 solar radii. The error bars represent the standard deviation of multiple rotation period and $v \sin(i)$ measurements. The figure demonstrates that the photometrically measured rotation periods are consistent with the spectroscopically derived rotation velocities.

3. The distribution of gyrochronology ages for the 30 NGC 6819 members. We have calculated the gyro ages for 21 stars with $(B - V)_0$ between 0.55 and 0.9 mag (masses between ~ 1.1 and 0.85 solar masses) using the model of ref. 11. Although this model is valid for stars with $(B - V)_0 > 0.47 \text{ mag}$, the NGC 6819 rotation sequence is poorly defined (by only two stars) for $0.47 \text{ mag} < (B - V)_0 < 0.55 \text{ mag}$.

For each of the 21 stars we have converted its $(B - V)_0$ colour into a value of the (global) convective turnover timescale, τ , using a numerical one-to-one transformation table³⁶, and linear interpolation as needed. The resulting τ values are associated with the measured rotation period values, P , and inserted into equation (32) in ref. 11, that is

$$t = (\tau/k_C) \ln(P/P_0) + (k_I/2\tau)(P^2 - P_0^2)$$

using an initial period of $P_0 = 1.1 \text{ d}$, as suggested in ref. 11. The dimensionless constants $k_C = 0.646 \text{ Myr d}^{-1}$ and $k_I = 452 \text{ d Myr}^{-1}$ listed there have been retained unmodified. This expression provides the age, t (in Myr), explicitly in terms of the independent variables, P and τ .

All of the resulting individual ages lie between 2 and 3 Gyr, as can be seen in the histogram in Extended Data Fig. 9, which is peaked between these values. The formal mean age is 2.49 Gyr ($1\sigma = 0.25 \text{ Gyr}$), and the median age is 2.43 Gyr. The gyrochronology age for NGC 6819 is therefore 2.49 Gyr with a standard error of 0.056 Gyr (2%, ignoring possible systematic errors in gyrochronology).

4. Is contaminating light from close neighbours a problem? NGC 6819 contains about 2,500 stars^{25,37}, is located 2.4 kiloparsecs ($\sim 7,800 \text{ light years}$) from the Sun^{24,37–39} and is only 13.5° above the Galactic plane. The cluster field is therefore densely populated with both cluster and foreground field stars. Accordingly, we must consider whether light from nearby stars can have 'leaked' into the photometric apertures used for our cluster targets. It would be difficult, if not impossible, to eliminate such contamination for all cluster stars with rotation period measurements. Instead, we take a qualitative approach to building a strong case against our overall result being unduly influenced by contaminating neighbours.

To accomplish this, we took advantage of our extensive stellar catalogue for the NGC 6819 field based on the deep, high-spatial-resolution CFHT/MegaCam images²⁵. For each of 43 cluster stars for which we initially detected periodic photometric variability, we listed the astrometric and photometric properties for all neighbouring stars within a 20 arcsec radius. We also used the Kepler Guest Observer

tool 'kepfeld' to extract the pixel mask images (PMI) for each target star and for each quarter of observations. The 'kepfeld' tool provides the coordinates, Kepler magnitudes and Kepler IDs for all stars in the Kepler Input Catalog (KIC) within the PMI. We searched for periodic variability in the light curves for all such neighbours observed by Kepler. Collectively, this information allowed us to study the angular separation, relative brightness and variability of neighbours within a ~ 5 pixel radius from each of the variable cluster stars. For the brightest neighbours, the photometric colour indices, effective temperatures and surface gravities derived from our star catalogue or given in the KIC, or both, provided additional guidance regarding their spectral type and evolutionary state.

From the 43 stars of interest, we rejected 13. Some of these stars were removed because the detected photometric variability was equal in period and in phase to variability detected in a close neighbour. Others were removed because of correlations between the presence and amplitude of periodic variability and the shape and size of the photometric aperture. Any such correlation suggests that the source of the variable signal originates outside the aperture and that a change in the aperture's size, shape and location results in more or less of the contaminating light being included. For some of the discarded stars the shape and size of the photometric aperture used in a given quarter was clearly affected by the signal from neighbouring stars. For others their location was significantly offset from the pixel with maximum counts inside the aperture, or, in the most severe cases, was entirely outside the aperture in one or more quarters.

Extended Data Fig. 7 shows the PMIs for three quarters (Q) for one accepted star and one quarter each for two of the rejected stars. Extended Data Fig. 7a displays the PMI for Q15, Q16 and Q17 for accepted star KIC 4938993 (green circle within optimal aperture). This ~ 1.2 solar-mass star ($V = 15.7$ mag, $(B - V)_0 = 0.5$ mag) has 13 neighbours within 20 arcsec. The brightest neighbour is 1 mag brighter than KIC 4938993 and is located 19.6 arcsec away and outside the PMI. All other neighbours are 4 arcsec or more distant and 2.8 mag or more fainter. None of these neighbours has a light curve available from the Kepler archive. KIC 4938993 dictates the shape, location and size of the optimal aperture and falls on the optimal-aperture pixel with maximum counts in all quarters. It is thus highly unlikely that the 11.89 d, ~ 30 mmag amplitude signal observed for this star originates in a neighbouring star. In Extended Data Fig. 7b the PMI for quarter 8 for rejected star KIC 5023712 is displayed. The KIC star near the left edge of the PMI is 3.2 mag brighter than KIC 5023712, and was not observed by Kepler. The star at (column, row) = (592.1, 853.8) is 0.43 mag fainter and it is closer to the optimal aperture of KIC 5023712 than the 2 pixel radius of the circular aperture that captures 95% of its signal. This star varies with the same period and three times the amplitude of KIC 5023712, suggesting that the periodic signal observed for KIC 5023712 originates in this fainter neighbour. Extended Data Fig. 7c shows the PMI for quarter 5 for rejected star KIC 5287900 and how a nearby star brighter by 4.7 mag causes the optimal aperture to shift off the target entirely.

Following this analysis, we were left with 30 stars for which we believe the periodic variability in their light curves reflects rotational modulation. Our confidence in the 30 rotation periods is bolstered by the comparison between their expected projected rotation velocities and their spectroscopically measured $v \sin(i)$ values (Extended Data Fig. 8). Furthermore, the narrow rotational sequence traced by the 30 stars in the NGC 6819 CPD (Fig. 2) is by itself strong evidence against significant contamination. Neighbouring stars are likely to have masses (colours) or ages, or both, different from those of our target stars, and significant contamination would broaden the observed narrow sequence.

5. Are tidal interactions with close companions a concern for gyrochronology?

The gyrochronology age of a cool star is derived under the assumption that its rotational evolution has not been influenced by external forces. Although this is the case for the vast majority of cool Galactic field stars, a small fraction of stars will

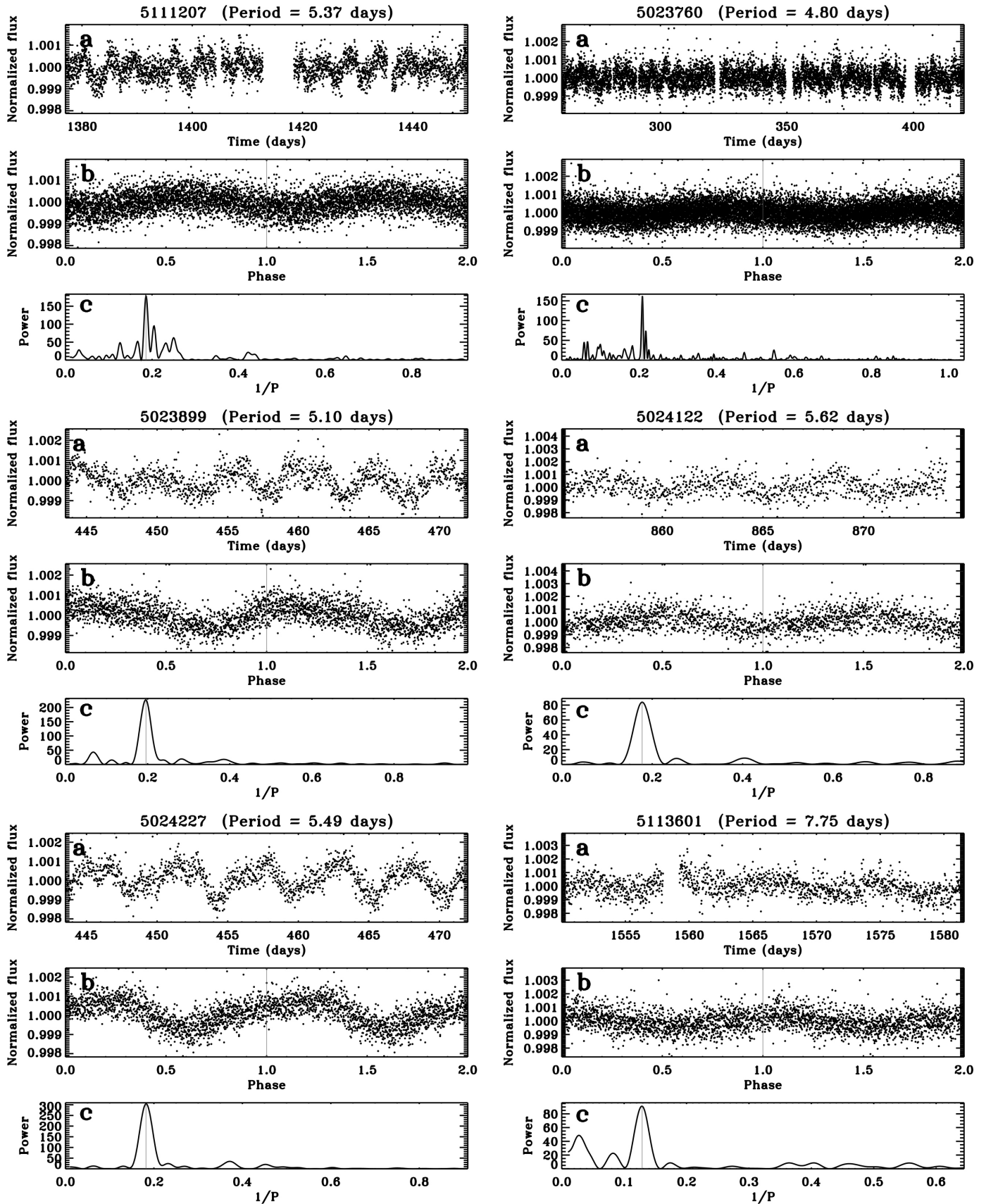
have a stellar or planetary companion close enough that tidal forces can drive an exchange of spin and orbital angular momentum between the two objects^{40,41}. We provide here an estimate of the fraction of cool Galactic field stars for which tidal interactions can potentially affect their rotation.

The tidal torque scales as the inverse of the binary semi-major axis to the sixth power, restricting significant tidal evolution for unevolved cool stars to the very closest systems. For stellar binaries with solar-type components, tidal synchronization will have a significant effect on the stellar rotation only in systems with orbital periods less than ~ 20 d (semi-major axis less than ~ 0.18 AU; ref. 32). Using the distribution of orbital periods for field binaries with solar-type components⁴², we find that binaries with periods of less than 20 d correspond to $\sim 5\%$ of that population. The same study⁴² estimates that less than half (46%) of all solar-type stars are in binaries, implying that only about $\sim 2.5\%$ of cool Galactic field stars have a stellar companion close enough for effective tidal interactions.

Short-period planetary companions can also tidally interact with their hosts, causing an increase in the stellar rotation period. The relevant class of planets here is the hot Jupiters, that is, planets with at least 10% of Jupiter's mass and orbits of less than 10 d (ref. 43). Such planets occur around solar-type field stars with a frequency of $\sim 1\%$ (ref. 43). Although the efficiency of tidal evolution in such star-planet systems is still uncertain on theoretical and observational grounds, it is probably only the hottest and most massive of the hot Jupiters that will have a significant effect on the rotational evolution of their host stars.

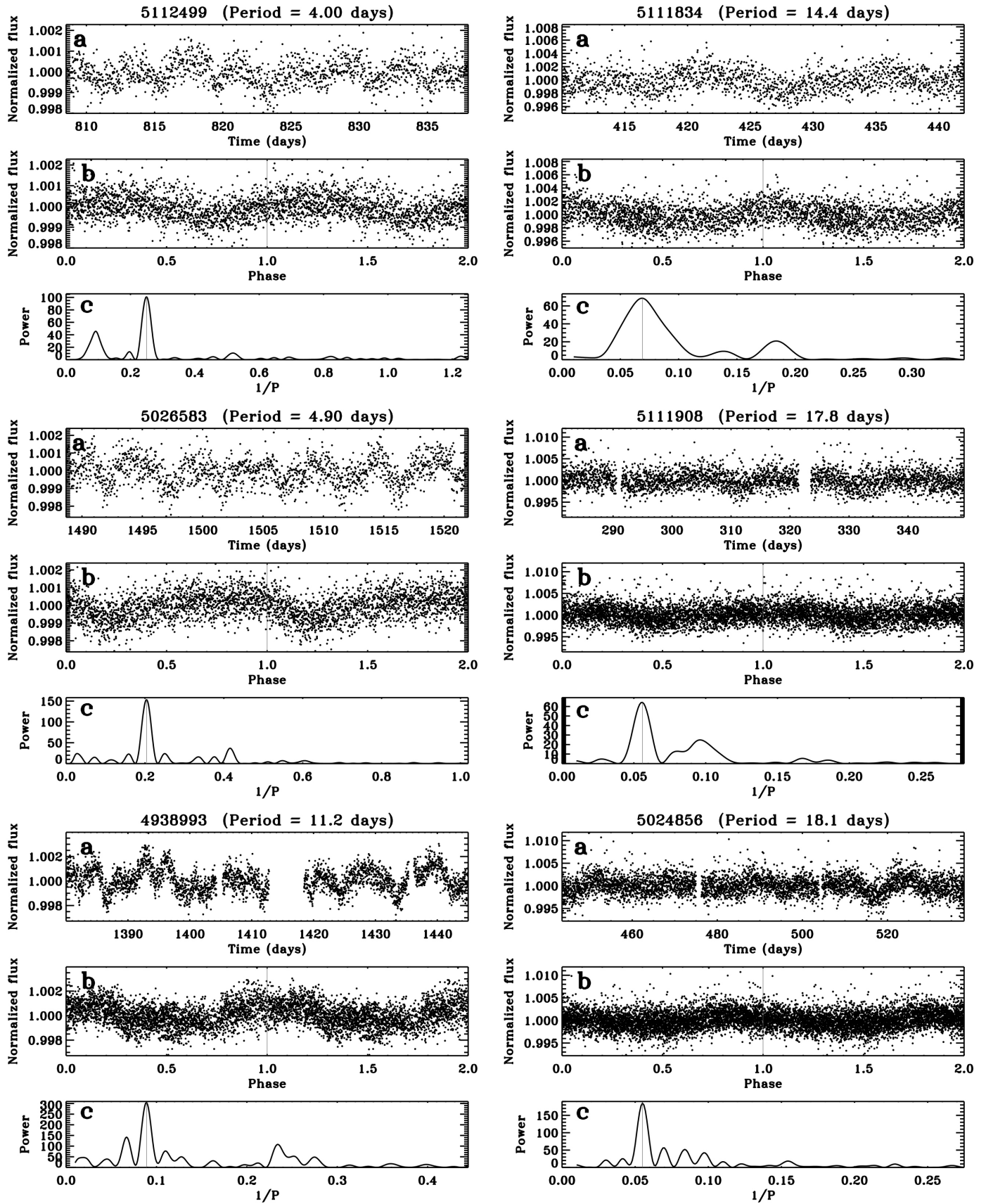
Assuming for simplicity that hot Jupiters do not occur in binaries with orbital periods shorter than 20 d, we estimate that at most 3% of cool stars have stellar or planetary companions that can affect their rotation through tidal interactions. Thus, tidal effects on rotation are not a concern for $\sim 97\%$ of cool Galactic field stars.

- Meibom, S. & Mathieu, R. D. A robust measure of tidal circularization in coeval binary populations: the solar-type spectroscopic binary population in the open cluster M35. *Astrophys. J.* **620**, 970–983 (2005).
- Meibom, S. *et al.* An observational study of tidal synchronization in solar-type binary stars in the open clusters M35 and M34. *Astrophys. J.* **653**, 621–635 (2006).
- Lasker, B. M. *et al.* The Second-Generation Guide Star Catalog: description and properties. *Astron. J.* **136**, 735–766 (2008).
- Smith, J. C. *et al.* Presearch Data Conditioning II - A Bayesian approach to systematic error correction. *Publ. Astron. Soc. Pacif.* **124**, 1000–1014 (2012).
- Stumpe, M. C. *et al.* Kepler Presearch Data Conditioning I - Architecture and algorithms for error correction in Kepler light curves. *Publ. Astron. Soc. Pacif.* **124**, 985–999 (2012).
- Barnes, S. A. & Kim, Y.-C. Angular momentum loss from cool stars: an empirical expression and connection to stellar activity. *Astrophys. J.* **721**, 675–685 (2010).
- Yang, S.-C. *et al.* WIYN Open Cluster Study LII: wide-field CCD photometry of the old open cluster NGC 6819. *Astron. J.* **762**, 3–20 (2013).
- Rosvick, J. M. & Vandenberg, D. A. BV photometry for the ~ 2.5 Gyr open cluster NGC 6819: more evidence for convective core overshooting on the main sequence. *Astron. J.* **115**, 1516–1523 (1998).
- Bragaglia, A. *et al.* Metal abundances of red clump stars in open clusters. I. NGC 6819. *Astron. J.* **121**, 327–336 (2001).
- Zahn, J.-P. Tidal friction in close binary stars. *Astron. Astrophys.* **57**, 383–394 (1977).
- Hut, P. Tidal evolution in close binary systems. *Astron. Astrophys.* **99**, 126–140 (1981).
- Raghavan, D. A survey of stellar families: multiplicity of solar type stars. *Astrophys. J.* **190** (suppl.), 1–42 (2010).
- Wright, J. T. The frequency of hot Jupiters orbiting nearby solar-type stars. *Astrophys. J.* **753**, 160–165 (2012).
- Brown, T. M. *et al.* Kepler Input Catalog: photometric calibration and stellar classification. *Astron. J.* **142**, 112–130 (2011).
- Anthony-Twarog, B. J. *et al.* A *uvby*CaH β analysis of the old open cluster, NGC 6819. *Astron. J.* **148**, 51–64 (2014).

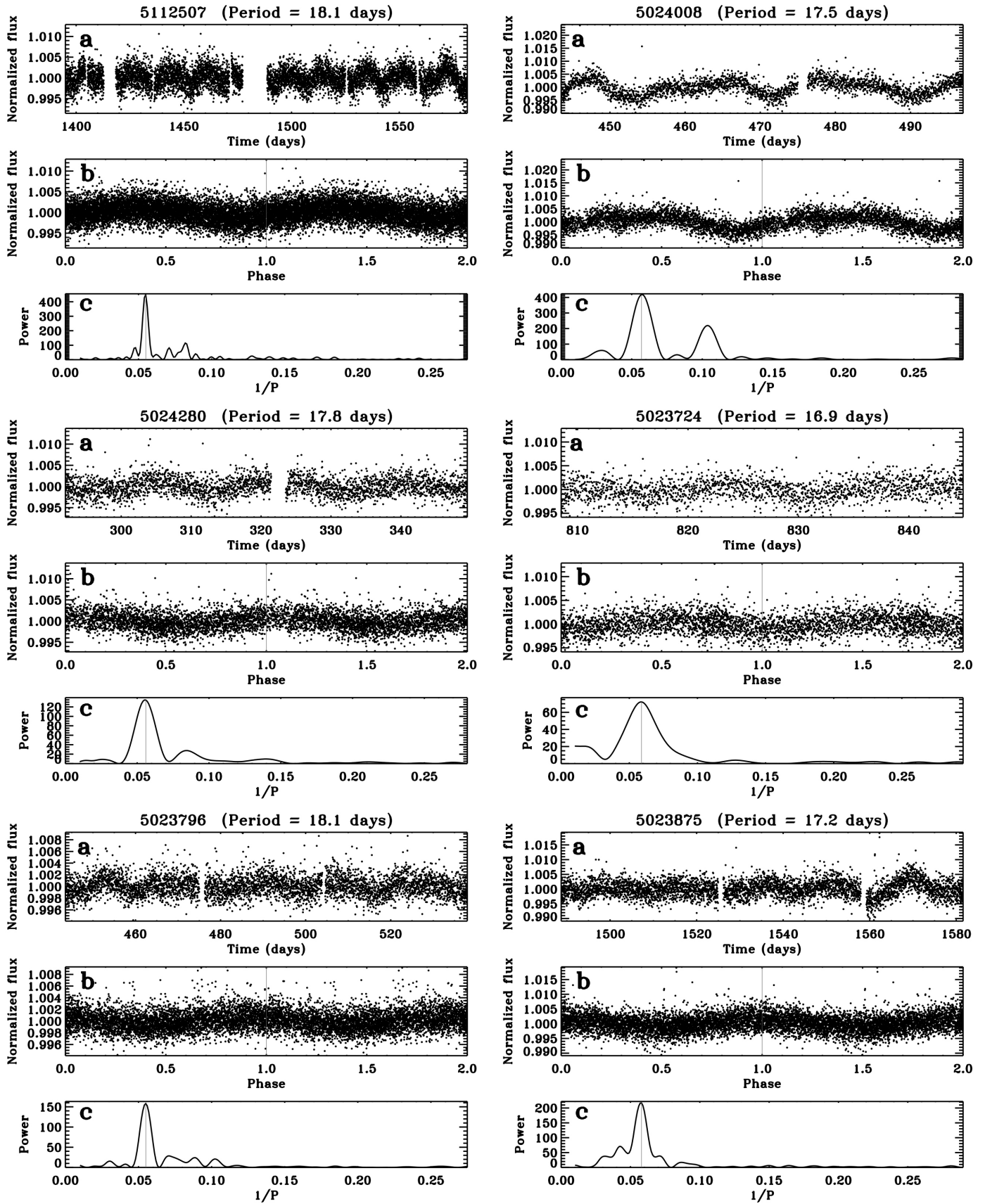


Extended Data Figure 1 | The light curves, phase-folded light curves and periodograms for the stars KIC 5111207, 5023899, 5024227, 5023760, 5024122 and 5113601. For each star we show a segment of the full Kepler light curve used for determining its rotation period (a), the corresponding

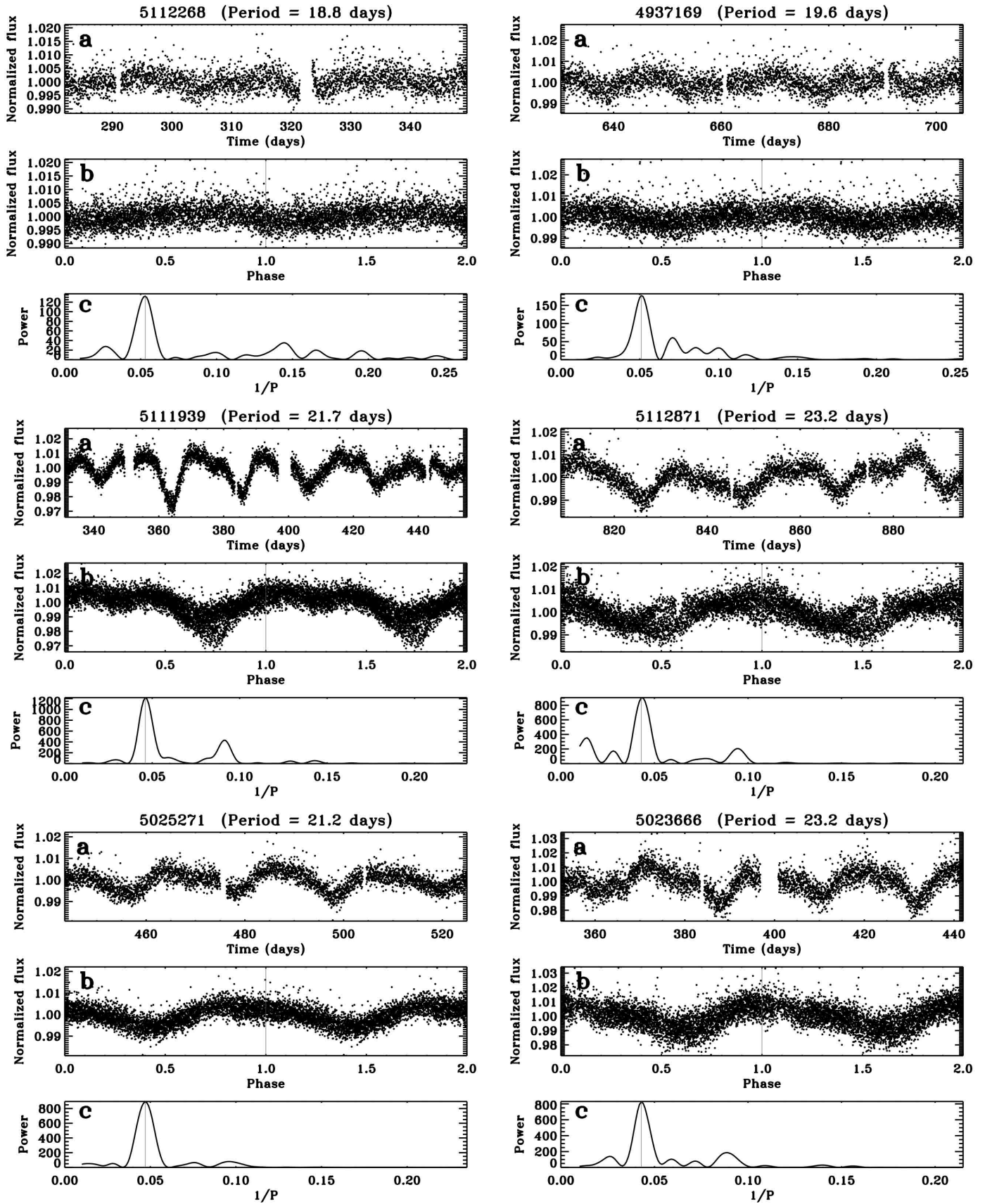
phase-folded light curve (b), and the periodogram (power as a function of rotation frequency, c). The KIC identification number⁴⁴ and the measured rotation period for each star are shown above the light curve segments.



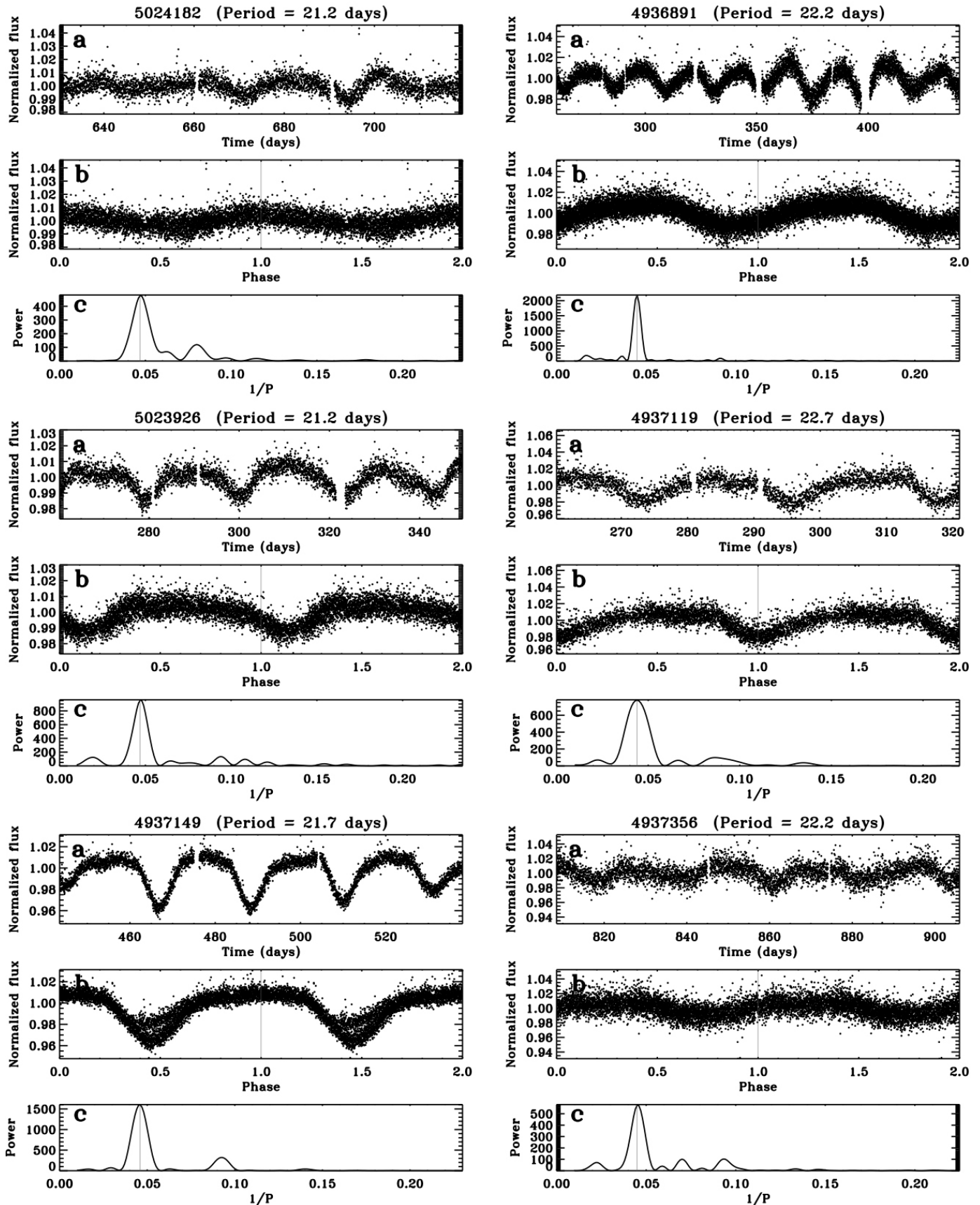
Extended Data Figure 2 | The light curves, phase-folded light curves and periodograms for the stars KIC 5112499, 5026583, 4938993, 5111834, 5111908 and 5024856. See Extended Data Fig. 1 for details.



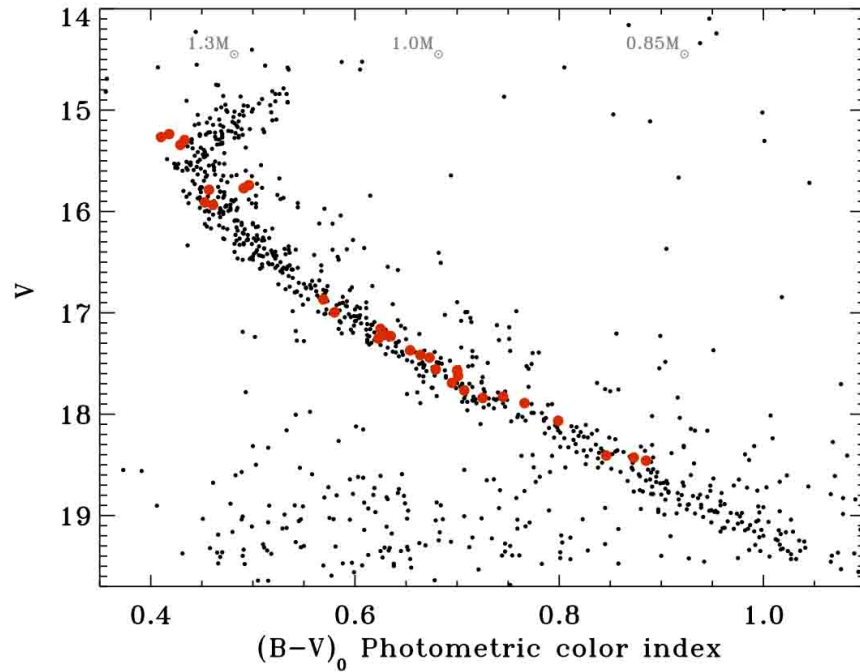
Extended Data Figure 3 | The light curves, phase-folded light curves, and periodograms for the stars KIC 5112507, 5024280, 5023796, 5024008, 5023724 and 5023875. See Extended Data Fig. 1 for details.



Extended Data Figure 4 | The light curves, phase-folded light curves and periodograms for the stars KIC 5112268, 5111939, 5025271, 4937169, 5112871 and 5023666. See Extended Data Fig. 1 for details.

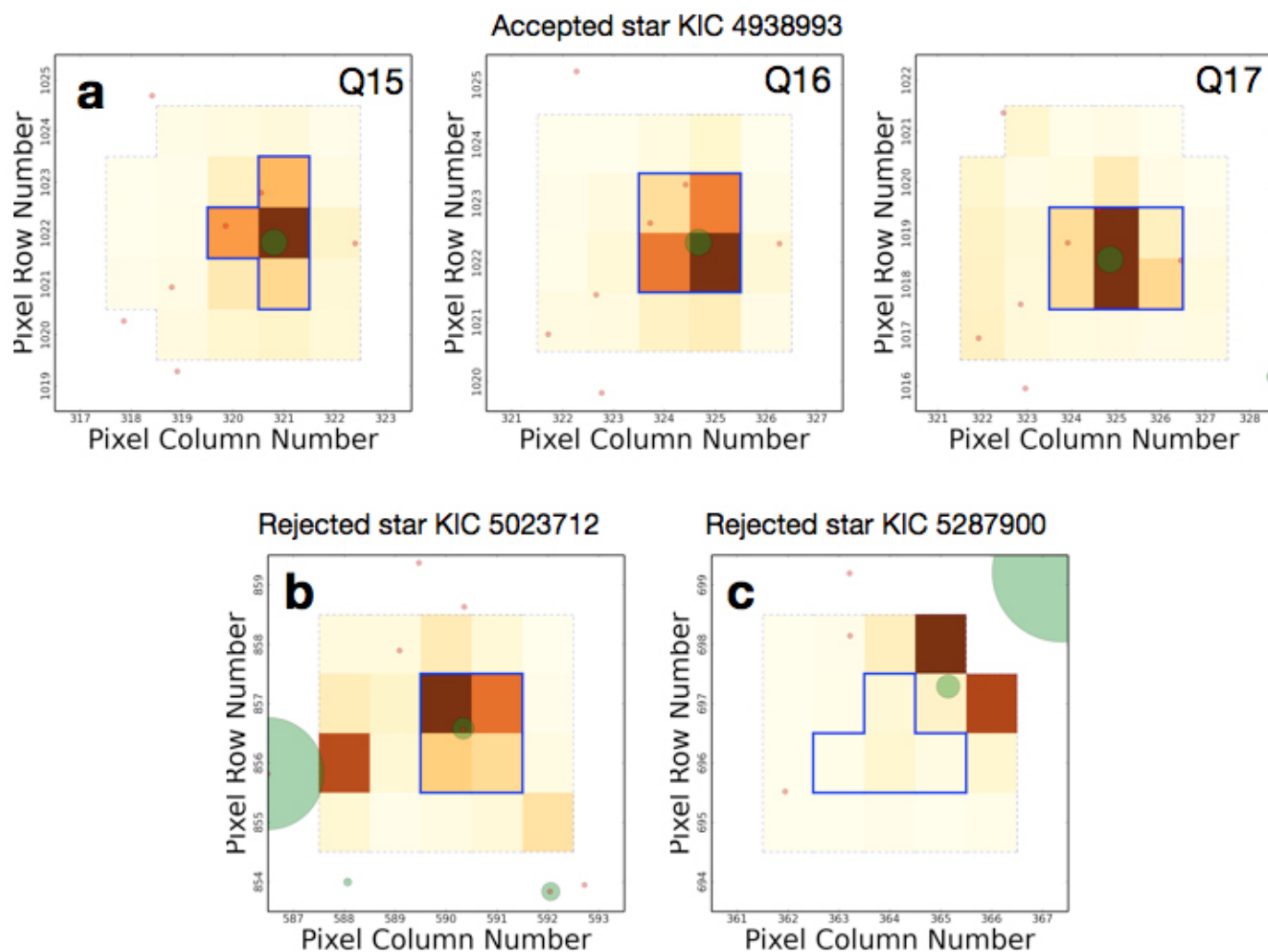


Extended Data Figure 5 | The light curves, phase-folded light curves and periodograms for the stars KIC 5024182, 5023926, 4937149, 4936891, 4937119 and 4937356. See Extended Data Fig. 1 for details.



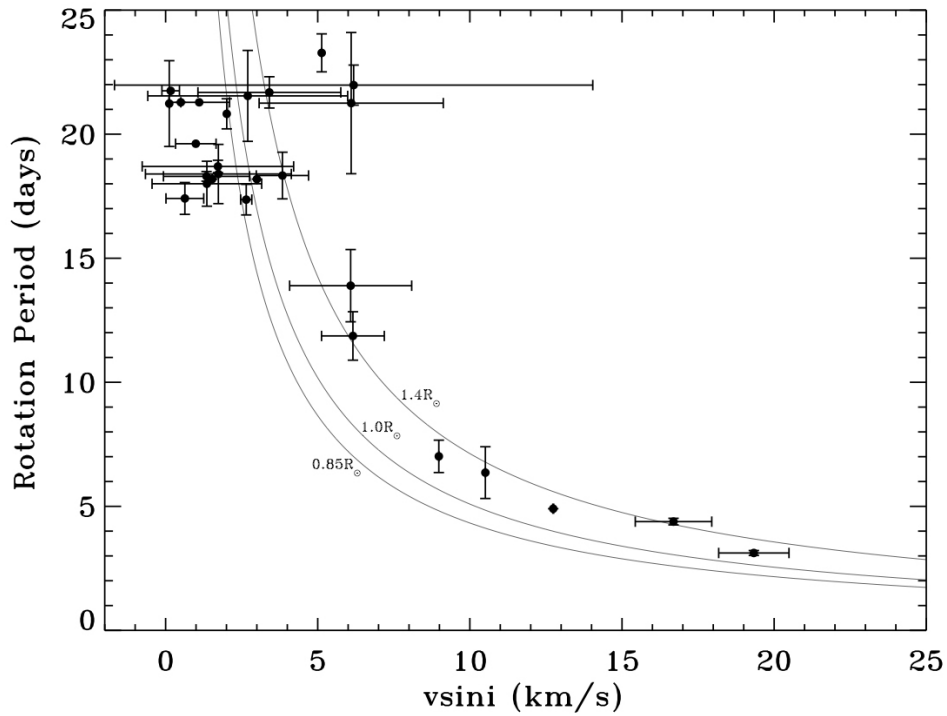
Extended Data Figure 6 | The NGC 6819 colour-magnitude diagram. The colour-magnitude diagram for stars identified as common proper-motion members of NGC 6819²⁵ and located within 5 arcmin of the cluster centre. The diagonal band tracing a tight relationship between the de-reddened photometric colour index, $(B - V)_0$, and brightness, V , represents the population of cluster members. The locations of the 30 stars with measured

rotation periods are marked with larger red circles. They all fall along this band and are thus photometric members of NGC 6819. Stellar masses in solar units are given along the top horizontal axis at the corresponding colours. The light from distant binary companions causes the two rotators near $(B - V)_0 = 0.5$ mag to fall above the cluster sequence.



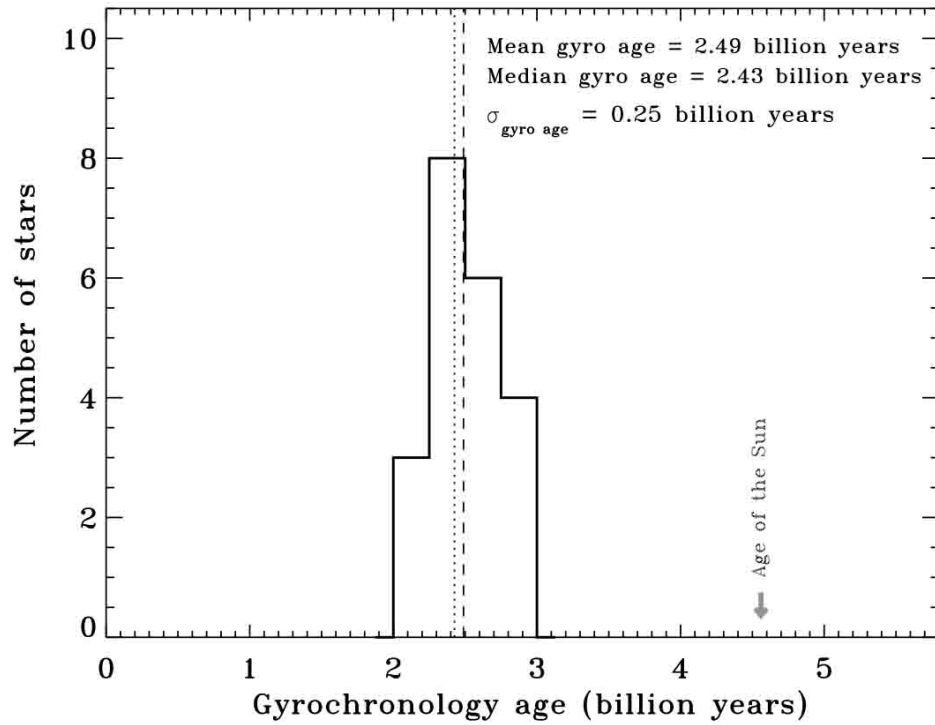
Extended Data Figure 7 | Pixel mask images for NGC 6819 members. Examples of pixel mask images (PMI) for the accepted star KIC 4938993 (**a**) and the rejected stars KIC 5023712 (**b**) and KIC 5287900 (**c**). Semi-transparent green circles mark the positions of KIC sources. Red dots correspond to the positions of fainter sources from deeper surveys within the

Kepler field. The solid blue line traces the optimal aperture (optimizing the signal-to-noise ratio for the target) defined for each target star in each quarter. The shape, size and location of the optimal aperture typically differed for the different quarters of observations.



Extended Data Figure 8 | A comparison of rotation periods and projected rotation velocities for stars in NGC 6819. Projected rotation velocities ($v \sin(i)$) plotted against the measured rotation periods for stars in NGC 6819. For comparison, three solid black curves show the expected relations between rotation period and $v \sin(i)$ for stars with respective radii of 0.85, 1.0 and

1.4 solar radii, observed at an inclination angle of their spin axes (i) of 90° . All stars plotted have single-lined spectra. The average rotation velocity resolution in the Hectochelle spectra is 7.38 km s^{-1} . The agreement between the expected and observed $v \sin(i)$ values for the measured rotation periods provides additional validation of our rotation period measurements.



Extended Data Figure 9 | The gyro age distribution for 21 cool dwarf members of NGC 6819. The gyrochronology ages for the 21 stars in the NGC 6819 CPD with $(B - V)_0$ colours in the range from 0.55 to 0.9 mag (masses between ~ 1.1 and 0.85 solar masses). The mean and median of the

distribution are 2.49 and 2.43 Gyr, respectively. The standard deviation for the 21 gyro ages is 0.25 Gyr, or 10% of the mean gyro age for the cluster. The standard error of the 2.49 Gyr mean is 0.056 Gyr, or $\sim 2\%$.

Extended Data Table 1 | Basic parameters and rotation period measurements for 30 members of NGC 6819

Kepler ID	Right Ascension	Declination	V	(B-V) ₀	nP	P _{Mean}	σ_P	P _{RV}	P _{μ}
	(hr min sec)	(° ' ")	(magn)	(magn)		(Days)	(Days)	(%)	(%)
5111207	19 40 09.01	40 12 16.7	15.27	0.41	7	5.42	1.05	92	61
5023899	19 40 55.90	40 10 04.7	15.24	0.42	6	4.81	0.44	92	99
5023760	19 40 49.33	40 09 07.0	15.30	0.43	6	4.78	0.10	75	99
5024227	19 41 09.49	40 11 41.6	15.34	0.43	3	5.06	0.39	88	99
5024122	19 41 05.50	40 08 28.4	15.91	0.45	2	6.36	1.04	93	96
5112499	19 41 13.91	40 15 30.3	15.94	0.46	3	4.44	0.92	93	99
5113601	19 42 00.18	40 12 19.2	15.79	0.46	5	7.01	0.65	92	99
5026583	19 42 55.25	40 09 48.0	15.77	0.49	1	4.90	-	91	36
4938993	19 42 55.85	40 00 19.8	15.74	0.50	3	11.87	0.98	81	62
5111834	19 40 43.86	40 15 18.4	16.87	0.57	7	13.89	1.46	92	97
5111908	19 40 43.86	40 12 53.2	17.00	0.58	2	17.41	0.64	93	97
5024856	19 41 30.61	40 08 42.8	17.26	0.62	1	18.19	-	76	95
5024280	19 41 11.63	40 08 34.1	17.22	0.63	3	17.36	0.62	93	96
5112507	19 41 14.09	40 17 34.7	17.16	0.63	1	18.19	-	91	97
5023796	19 40 51.10	40 09 45.1	17.23	0.64	3	18.30	0.19	67	96
5024008	19 41 00.51	40 10 29.4	17.37	0.65	2	18.40	1.19	93	96
5023724	19 40 47.50	40 07 40.3	17.42	0.66	3	18.00	0.91	91	92
5023875	19 40 54.82	40 09 20.5	17.44	0.67	3	18.33	0.94	77	95
5112268	19 41 04.25	40 14 23.2	17.56	0.68	2	18.70	0.25	91	94
4937169	19 41 22.61	40 05 44.6	17.62	0.70	1	19.62	-	93	95
5025271	19 41 47.15	40 09 30.3	17.57	0.70	2	21.29	0.10	89	93
5111939	19 40 49.63	40 12 05.2	17.70	0.70	1	21.75	-	92	90
5112871	19 41 26.64	40 16 01.4	17.77	0.71	2	21.25	2.85	78	86
5023666	19 40 44.23	40 10 11.6	17.84	0.73	3	21.54	1.83	92	88
5024182	19 41 07.57	40 08 50.3	17.83	0.75	1	21.29	-	90	92
5023926	19 40 56.89	40 06 10.1	17.89	0.77	4	20.82	0.61	93	91
4937149	19 41 21.54	40 05 21.1	18.07	0.80	6	21.68	0.63	92	88
4936891	19 41 07.75	40 05 55.0	18.41	0.85	7	21.98	0.81	92	82
4937119	19 41 19.90	40 05 39.7	18.43	0.87	2	23.28	0.77	89	59
4937356	19 41 33.09	40 05 12.4	18.46	0.89	3	21.23	1.73	93	78

The Kepler ID is from the KIC⁴⁴. Right ascension and declination are equinox J2000²⁵. The stellar brightness, V , and colour index, $(B - V)_0$, are based on CCD photometry using the CFH12K mosaic CCD on the 3.6 m CFHT²⁴. The $B - V$ colours were de-reddened using a value for the colour excess $E_{(B-V)}$ of 0.15 (refs 38, 39, 45). For stars with more than one rotation period measurement ($nP > 1$), the P_{Mean} and σ_P columns list the mean period and the standard deviation of the multiple measurements.

Michelson–Morley analogue for electrons using trapped ions to test Lorentz symmetry

T. Pruttivarasin^{1,2}, M. Ramm¹, S. G. Porsev^{3,4}, I. I. Tupitsyn⁵, M. S. Safronova^{3,6}, M. A. Hohensee^{1,7} & H. Häffner¹

All evidence so far suggests that the absolute spatial orientation of an experiment never affects its outcome. This is reflected in the standard model of particle physics by requiring all particles and fields to be invariant under Lorentz transformations. The best-known tests of this important cornerstone of physics are Michelson–Morley-type experiments verifying the isotropy of the speed of light^{1–3}. For matter, Hughes–Drever-type experiments^{4–11} test whether the kinetic energy of particles is independent of the direction of their velocity, that is, whether their dispersion relations are isotropic. To provide more guidance for physics beyond the standard model, refined experimental verifications of Lorentz symmetry are desirable. Here we search for violation of Lorentz symmetry for electrons by performing an electronic analogue of a Michelson–Morley experiment. We split an electron wave packet bound inside a calcium ion into two parts with different orientations and recombine them after a time evolution of 95 milliseconds. As the Earth rotates, the absolute spatial orientation of the two parts of the wave packet changes, and anisotropies in the electron dispersion will modify the phase of the interference signal. To remove noise, we prepare a pair of calcium ions in a superposition of two decoherence-free states, thereby rejecting magnetic field fluctuations common to both ions¹². After a 23-hour measurement, we find a limit of $h \times 11$ millihertz (h is Planck's constant) on the energy variations, verifying the isotropy of the electron's dispersion relation at the level of one part in 10^{18} , a 100-fold improvement on previous work⁹. Alternatively, we can interpret our result as testing the rotational invariance of the Coulomb potential. Assuming that Lorentz symmetry holds for electrons and that the photon dispersion relation governs the Coulomb force, we obtain a fivefold-improved limit on anisotropies in the speed of light^{2,3}. Our result probes Lorentz symmetry violation at levels comparable to the ratio between the electroweak and Planck energy scales¹³. Our experiment demonstrates the potential of quantum information techniques in the search for physics beyond the standard model.

Invariance under Lorentz transformations is a key feature of the standard model, and as such is fundamental to nearly every aspect of modern physics. Nevertheless, this symmetry may be measurably violated, for example, as a result of spontaneous symmetry breaking in quantum fields with dynamics at experimentally inaccessible energy scales not explicitly treated by the standard model¹⁴. Some theories that unify gravitation and the standard model assert that Lorentz symmetry is valid only at large length scales^{15,16}. A natural estimate of the fractional shift of electron dispersion relations due to Lorentz violation at the Planck scale is given by the ratio between the electroweak and Planck energy scales, that is, $\sim 2 \times 10^{-17}$ (ref. 13). Other models suggest that large Lorentz violation at the Planck scale is suppressed by supersymmetry¹⁷. In such scenarios, the constraints on Lorentz violation for neutrons⁶ can be used to set an upper bound of order 100 TeV on the supersymmetric energy scale¹⁸. Therefore, precision tests of Lorentz symmetry complement

direct probes of high-energy physics being carried out at the Large Hadron Collider.

We analyse Lorentz violation in the context of a phenomenological framework known as the standard model extension^{19,20} (SME). The SME is an effective field theory that augments the standard model Lagrangian with every possible combination of the standard model fields that is not term-by-term Lorentz invariant, while maintaining gauge invariance, energy–momentum conservation, and Lorentz invariance of the total action^{19,20}. The SME can be used to describe the low-energy limit of many different theories which predict Lorentz violation, and includes the standard model as a limiting case. The SME thus provides a comprehensive framework for quantifying a wide range of Lorentz-violating effects, and is a flexible tool for consistently evaluating a wide variety of experiments²¹.

The SME allows for Lorentz violation for all particles separately. However, to verify a particle's Lorentz symmetry, it must be compared with a reference system because only differences in their behaviours under Lorentz transformation are observable²⁰. For instance, typical interpretations of Michelson–Morley experiments testing Lorentz violation of photons assume that the lengths of the interferometer arms are invariant under rotations. Because the lengths of interatomic bonds depend on the electron dispersion relation^{22,23}, those interpretations can be said to assume that Lorentz symmetry for electrons (and nuclei making up the interferometer arms) holds unless a second distinct reference system is used²³. For our experiment, it seems more natural to use light as a reference and assume that photons obey Lorentz symmetry. However, it is important to keep in mind that an experimental signature of the Lorentz violation considered here can equally be attributed to Lorentz violation of electrons as well as to that of photons, which would manifest itself as an asymmetry of the photon-mediated Coulomb potential (Methods). Thus, we take the most general view, namely that we measure the difference between the electron and photon anisotropies.

We take this view by choosing a coordinate system in which a hypothetical Lorentz violation in light manifests itself in the electronic Lagrangian (Methods). We obtain the modified electronic quantum electrodynamics Lagrangian

$$\mathcal{L} = \frac{1}{2} i \bar{\psi} \left(\gamma_\nu + c'_{\mu\nu} \gamma^\mu \right) \bar{D}^\nu \psi - \bar{\psi} m_e \psi \quad (1)$$

where m_e is the electron mass, ψ is a Dirac spinor, γ^μ are the Dirac matrices and $\bar{\psi} \bar{D}^\nu \psi \equiv \bar{\psi} D^\nu \psi - \psi D^\nu \bar{\psi}$ with D^ν being the covariant derivative. The effect due to Lorentz violation is described by the tensor $c'_{\mu\nu} = c_{\mu\nu} + k_{\mu\nu}/2$, which contains Lorentz-violation parameters from both the electron ($c_{\mu\nu}$) and the photon ($k_{\mu\nu}$) sectors^{19,20}. Because $c'_{\mu\nu}$ is frame dependent, we uniquely specify its value in the Sun-centred, celestial-equatorial frame (SCCEF), that is, the Sun's rest frame. Time-dependent Lorentz transformations due to the Earth's motion transform $c'_{\mu\nu}$ in the SCCEF to the time-dependent values in the local laboratory frame

¹Department of Physics, University of California, Berkeley, California 94720, USA. ²Quantum Metrology Laboratory, RIKEN, Wako, Saitama 351-0198, Japan. ³Department of Physics and Astronomy, University of Delaware, Newark, Delaware 19716, USA. ⁴Petersburg Nuclear Physics Institute, Gatchina, Leningrad District 188300, Russia. ⁵Department of Physics, St Petersburg State University, Ulianovskaya 1, Petrodvorets, St Petersburg 198504, Russia. ⁶Joint Quantum Institute, National Institute of Standards and Technology and the University of Maryland, College Park, Maryland 20742, USA. ⁷Lawrence Livermore National Laboratory, Livermore, California 94550, USA.

on the Earth. Hence, the contribution of $c'_{\mu\nu}$ to any laboratory-frame observable will vary in time.

For us, the important consequence of electronic Lorentz violation is the dependence of an electron's energy on the direction of its momentum. For an atomically bound electron with momentum \mathbf{p} , the Lagrangian in equation (1) results in a small energy shift that depends on the direction of the electron's momentum and is described by the effective Hamiltonian²⁴

$$\delta\mathcal{H} = -C_0^{(2)} \frac{(\mathbf{p}^2 - 3p_z^2)}{6m_e} \quad (2)$$

where $C_0^{(2)}$ contains elements in $c'_{\mu\nu}$ in the laboratory frame and p_z is the component of the electron's momentum along the quantization axis, which is fixed in the laboratory. The energy shift depends on how the total momentum \mathbf{p} is distributed among the three spatial components. As the Earth rotates, $C_0^{(2)}$ varies in time, resulting in a time variation of the electron's energy correlated with the Earth's motion.

To probe Lorentz violation, we perform the electronic analogue of a Michelson–Morley experiment by interfering atomic states with anisotropic electron momentum distributions aligned along different directions, such as available in the $^2D_{5/2}$ manifold of $^{40}\text{Ca}^+$. We trap a pair of $^{40}\text{Ca}^+$ ions with an ion–ion separation of $\sim 16\ \mu\text{m}$ in a linear Paul trap, and define the quantization axis by applying a static magnetic field of 3.930 G vertically. The direction of this magnetic field changes with respect to the Sun as the Earth rotates, resulting in a rotation of our interferometer (Fig. 1).

We calculate the Lorentz-violation-induced hypothetical energy shift of $^{40}\text{Ca}^+$ in the $^2D_{5/2}$ manifold according to equation (2) (expressed here in hertz):

$$\frac{\Delta E}{h} = [(2.16 \times 10^{15}) - (7.42 \times 10^{14})m_J]C_0^{(2)}$$

where m_J is the magnetic quantum number (Methods). To obtain maximum sensitivity to Lorentz violation, we monitor the energy difference between the states $|\pm 5/2\rangle \equiv |^2D_{5/2}; m_J = \pm 5/2\rangle$ and $|\pm 1/2\rangle \equiv |^2D_{5/2}; m_J = \pm 1/2\rangle$ using a Ramsey-type interferometric scheme. To reject magnetic field noise, which is the main source of decoherence, we create a product state $|\Psi^P\rangle = (1/2)(|-1/2\rangle + |-5/2\rangle) \otimes (|+1/2\rangle + |+5/2\rangle)$ by applying to both ions a series of $\pi/2$ - and π -pulses on

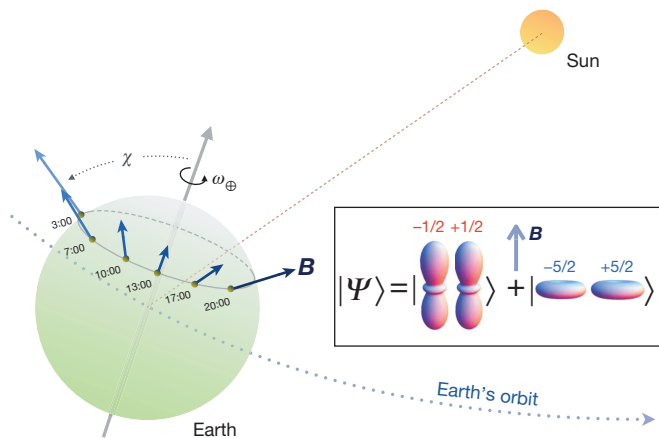


Figure 1 | Rotation of the quantization axis of the experiment with respect to the Sun as the Earth rotates. We apply a magnetic field (B) of 3.930 G vertically in the laboratory frame to define the quantization axis of the experiment. As the Earth rotates with an angular frequency given by $\omega_{\oplus} = 2\pi/(23.93\ \text{h})$, the orientation of the quantization axis and, consequently, that of the electron wave packet (as shown in the inset in terms of probability envelopes) changes with respect to the Sun's rest frame (positions at various times UTC are illustrated). The angle $\chi \approx 52.1^\circ$ is the colatitude of the experiment.

the $S_{1/2}$ – $D_{5/2}$ transition. Under common noise induced by a fluctuating magnetic field, the product state rapidly dephases to a mixed state that contains a decoherence-free entangled state $|\Psi^R\rangle \equiv (1/\sqrt{2})(|-5/2, +5/2\rangle + |-1/2, +1/2\rangle)$ with 50% probability²⁵. This entangled state evolves freely in time according to

$$|\Psi^R(t)\rangle = \frac{1}{\sqrt{2}}(|-5/2, +5/2\rangle + e^{i(\Delta E_R t/\hbar + \phi_R)}|-1/2, +1/2\rangle)$$

where ΔE_R is the energy difference between the states $|-5/2, +5/2\rangle$ and $|-1/2, +1/2\rangle$, ϕ_R is a phase offset and \hbar is Planck's constant divided by 2π . The remaining components of the mixed state, which are the states $|-5/2, +1/2\rangle$ and $|-1/2, +5/2\rangle$, each with 25% probability, are time independent.

In Fig. 2, we illustrate the dynamics of the state $|\Psi^R\rangle$. By expressing the state in the even–odd parity basis, $|\pm\rangle = (1/\sqrt{2})(|-5/2, +5/2\rangle \pm |-1/2, +1/2\rangle)$, the time-dependent state can be written as

$$|\Psi^R(t)\rangle = \frac{1}{\sqrt{2}}\left(\left(1 + e^{i(\Delta E_R t/\hbar + \phi_R)}\right)|+\rangle + \left(1 - e^{i(\Delta E_R t/\hbar + \phi_R)}\right)|-\rangle\right) \quad (3)$$

We interpret the trajectory of $|\Psi^R(t)\rangle$ to be along the equator of the Bloch sphere as shown in Fig. 2b. The state $|\Psi^R(t)\rangle$ oscillates back and forth between the states $|+\rangle$ and $|-\rangle$ with frequency $f_R = \Delta E_R/\hbar$. To read out the ion state in the $|\pm\rangle$ basis, we apply to both ions a series of π - and $\pi/2$ -pulses on the $S_{1/2}$ – $D_{5/2}$ transition, followed by an electron-shelving readout scheme¹². The difference between the probabilities

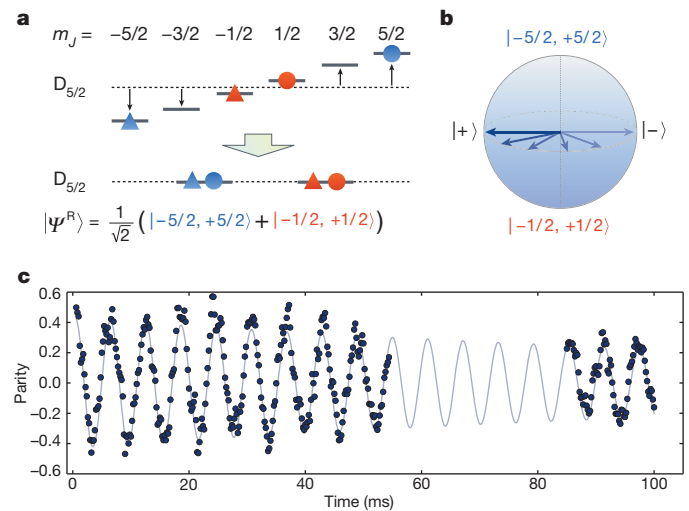


Figure 2 | Oscillation of the decoherence-free state. **a**, A combination of different magnetic sublevels of the first (circles) and second (triangles) $^{40}\text{Ca}^+$ ions in the $^2D_{5/2}$ manifold forms a decoherence-free state $|\Psi^R\rangle \equiv (1/\sqrt{2})(|-5/2, +5/2\rangle + |-1/2, +1/2\rangle)$. Blue and red colours indicate pairing of the single-ion states in each component of $|\Psi^R\rangle$. **b**, Time evolution of the state $|\Psi^R(t)\rangle$ represented by a trajectory on a Bloch sphere with poles given by $|-5/2, +5/2\rangle$ and $|-1/2, +1/2\rangle$. (We neglected contributions from the states $|-5/2, +1/2\rangle$ and $|-1/2, +5/2\rangle$, which have no phase coherence.) The state $|\Psi^R(t)\rangle$ oscillates back and forth between the even–odd parity basis states, $|\pm\rangle$, as given in equation (3). **c**, Oscillation of a product state containing an entangled state $|\Psi^R\rangle$ with 50% probability. Each data point is derived from 200 repetitions of the Ramsey-type experimental cycle shown in Fig. 3a. The error bars (no larger than data symbols, and omitted to simplify figure) are obtained from requiring that the fit to the Ramsey fringe function (grey solid line) gives $\sqrt{\chi^2_{\text{reduced}}} = 1$ and assuming that the data are normally distributed. The fit yields an oscillation frequency of $164.9 \pm 0.1\ \text{Hz}$ and a decay constant of $155 \pm 17\ \text{ms}$, which is substantially smaller than the value expected from the lifetime of the $^2D_{5/2}$ state of $^{40}\text{Ca}^+$. We attribute the loss of coherence to the heating rate of the ion trap, of $\sim 0.2\ \text{quanta ms}^{-1}$, which degrades the quality of the analysis pulses for long Ramsey interrogation times. To save measurement time, data was not taken in the ~ 60 – $80\ \text{ms}$ interval.

P_+ and P_- of the ions being in the states $|+\rangle$ and $|-\rangle$, respectively, yields an oscillating signal given by $P = P_+ - P_- = \cos(\Delta E_R t / \hbar + \phi_R)$ (Fig. 2c).

We are interested in the variations in the energy difference between the $|\pm 5/2, \mp 5/2\rangle$ and $|\pm 1/2, \mp 1/2\rangle$ states due to Lorentz violation. However, the energy difference is also affected by linear Zeeman shifts from a residual magnetic field gradient, quadratic Zeeman shifts, electric quadrupole shifts from an electric field gradient, and a.c. Stark shifts from oscillating trapping fields^{26,27}. The contributions from the magnetic field gradient, which are of order 100 Hz, have opposite signs for the state $|\Psi^R\rangle$ and its mirrored counterpart, $|\Psi^L\rangle \equiv (1/\sqrt{2})(|+5/2, -5/2\rangle + |+1/2, -1/2\rangle)$. We can subtract out the contribution from the magnetic field gradient to the oscillation signal by taking the average frequency $\bar{f} = (f_R + f_L)/2$, where f_R and f_L are the oscillation frequencies of state $|\Psi^R\rangle$ and $|\Psi^L\rangle$, respectively (Extended Data Fig. 1). The remaining effects (except for Lorentz violation) are energy shifts of order of only a few hertz and are also directly related to external electromagnetic fields in the proximity of the ions. We expect these fields to

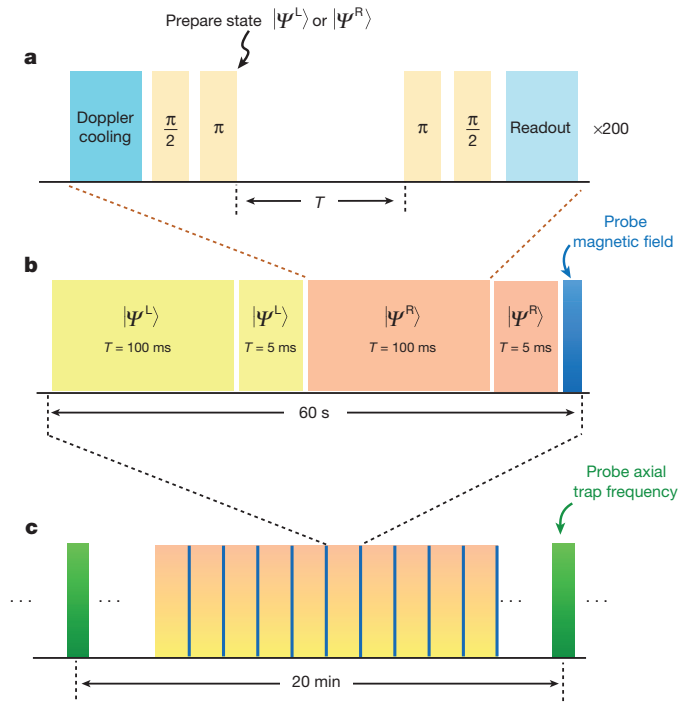


Figure 3 | Outline of the experimental scheme. **a**, The building blocks of our experiment are Ramsey-type interferometric sequences. In each measurement cycle, we first perform Doppler cooling and optical pumping of the ions. Then a series of $\pi/2$ - and π -pulses on the $S_{1/2}$ - $D_{5/2}$ transition prepare the ions in a product state that dephases into a mixed state within 1 ms. This state contains an entangled state $|\Psi^{L,R}\rangle \equiv (1/\sqrt{2})(|+5/2, \mp 5/2\rangle + |\pm 1/2, \mp 1/2\rangle)$ with 50% probability. Afterwards, the mixed state evolves freely for Ramsey duration T , before another series of π - and $\pi/2$ -pulses, together with an electron-shelving readout sequence, allows us to read out the state of the ions in the even-odd parity basis. This measurement cycle is repeated 200 times each for $|\Psi^L\rangle$ and $|\Psi^R\rangle$. **b**, To correct for phase drifts in the preparation of $|\Psi^{L,R}\rangle$, we measure the difference in the oscillation signal between Ramsey durations of 100 and 5 ms. We then correct for the contribution of the magnetic field gradient by taking the average of the oscillation signals measured with states $|\Psi^L\rangle$ and $|\Psi^R\rangle$. At the end of this measurement block, we measure the magnetic field by performing spectroscopy on the $S_{1/2}$ - $D_{5/2}$ transition to correct for the quadratic Zeeman effect. Each grey data point in Fig. 4a is a result from one of these measurement blocks. **c**, We continuously repeat the measurement block during the course of the 23 h-long measurement. To correct for the electric quadrupole shift caused by the electric field gradient, we measure the axial trap frequency by performing spectroscopy on the $S_{1/2}$ - $D_{5/2}$ transition.

be stable on the 10^{-3} level over the course of a day, and the associated variations are on the level of a few millihertz and below. Moreover, we independently measure these fields using the ions themselves as a probe (Methods).

We measured the energy difference between the states $|\pm 5/2, \mp 5/2\rangle$ and $|\pm 1/2, \mp 1/2\rangle$ of $^{40}\text{Ca}^+$ for 23 h starting from 3:00 coordinated universal time (UTC) on 19th April 2014, by monitoring the oscillation signal of the ions with an effective Ramsey duration of 95 ms (Methods). At the same time, we monitored the magnetic field and the electric field gradient using the ions themselves as a probe (Fig. 3). We then used the measured values of the magnetic field and electric field gradient to correct for the quadratic Zeeman and electric quadrupole shifts. The resulting 23 h frequency measurement is shown in Fig. 4. With 23 h of averaging, we reach a sensitivity of the oscillation frequency of 11 mHz, limited by statistical uncertainties due to short-term fluctuations. We then attribute any residual variation in the energy correlated with the Earth's rotation to Lorentz violation.

Lorentz transformations of ζ'_{uv} from the SCCEF to the laboratory frame results in the time-dependent energy shift due to Lorentz violation given by

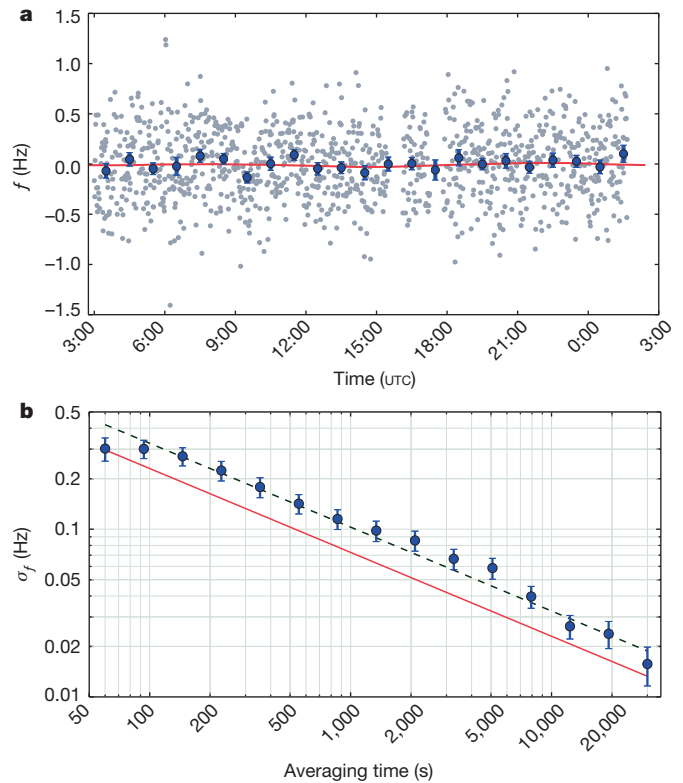


Figure 4 | Frequency measurements for $^{40}\text{Ca}^+$. **a**, The grey data points represent frequency measurements of $^{40}\text{Ca}^+$ taken after each measurement block as shown in Fig. 3b, with contributions from the quadratic Zeeman shifts and electric quadrupole shifts subtracted out. (Gaps in the data points are due to a failure of the laser frequency stabilization.) We started the measurement at 3:00 UTC on 19 April 2014, and continued for 23 h. The blue points are obtained by binning of data from 60 min time intervals. The error bars represent 1 s.d. of the data points within the bin, where we scale the error by $\sqrt{\chi^2_{\text{reduced}}}/N = 1.3$ (obtained from the fit of the binned data to the model in equation (4)). **b**, The blue points show the Allan deviation of the frequency measurement, σ_f , calculated from the unbinned data with error bars representing 1 s.d. The red solid line is the estimated quantum projection noise. The green dashed line is a fit to the data, showing a sensitivity to the ions' energy variation of $\sigma_t = 3.3 \text{ Hz}/\sqrt{\tau}$, where τ is the averaging time. The steady downward trend indicates that we are still limited by statistical fluctuations rather than by correlated noise or systematics over the course of the measurement.

Table 1 | Limits on differential electron–photon Lorentz-violation parameters $c'_{\mu\nu} = c_{\mu\nu} + k_{\mu\nu}/2$

Parameters ($c'_{\mu\nu} = c_{\mu\nu} + k_{\mu\nu}/2$)	New limits	Existing limits	
		$c_{\mu\nu}$ (electrons)	$k_{\mu\nu}/2$ (photons)
$-0.16c'_{XY} + 0.33c'_{XZ} - 0.92c'_{YZ} - 0.16c'_{YZ}$	$0.1 \pm 1.0 \times 10^{-18}$	$-0.9 \pm 1.0 \times 10^{-16}$	$-2.5 \pm 3.5 \times 10^{-18}$
$-0.04c'_{XY} - 0.32c'_{XZ} - 0.35c'_{YZ} + 0.88c'_{YZ}$	$2.4 \pm 7.4 \times 10^{-19}$	$-0.9 \pm 6.5 \times 10^{-17}$	$-5.2 \pm 3.6 \times 10^{-18}$
$0.29c'_{XY} - 0.38c'_{XZ} - 0.73c'_{YZ} - 0.48c'_{YZ}$	$5.9 \pm 9.5 \times 10^{-19}$	$-8.1 \pm 9.5 \times 10^{-17}$	$-0.6 \pm 3.8 \times 10^{-18}$
$-0.31c'_{XY} - 0.65c'_{XZ} + 0.07c'_{YZ} - 0.69c'_{YZ}$	$0.7 \pm 1.2 \times 10^{-18}$	$-2.9 \pm 6.5 \times 10^{-17}$	$-2.6 \pm 3.8 \times 10^{-18}$

Fitting our frequency measurements to the model in equation (4) yields the limits on Lorentz-violation parameters $c'_{\mu\nu}$ in the SCCEF. All uncertainties for the uncorrelated combinations of $c'_{\mu\nu}$ are 1 s.d. from the fit conservatively scaled with $\langle \chi^2_{\text{reduced}} \rangle = 1.3$. We improve the bounds from ref. 9 on the electron dispersion by up to two orders of magnitude. Alternatively, we can work in coordinates such that the electron dispersion is isotropic. We then improve by up to five times on the existing limits for the isotropy of the speed of light set by a modern version of the classic Michelson–Morley experiment in ref. 2 (Methods). Note that the work in ref. 9 assumed that $k_{\mu\nu} = 0$ whereas that in ref. 2 assumed that $c_{\mu\nu} = 0$. We use the notation $c'_{XY} = c'_{XX} - c'_{YY}$.

$$\frac{\Delta E}{h} = A \cos(\omega_{\oplus} T) + B \sin(\omega_{\oplus} T) + C \cos(2\omega_{\oplus} T) + D \sin(2\omega_{\oplus} T) \quad (4)$$

where $\omega_{\oplus} = 2\pi/(23.93 \text{ h})$ is the sidereal angular frequency of the Earth's rotation, T is the time since the vernal equinox of 2014, and A, B, C and D are parameters related to $c'_{\mu\nu}$ in the SCCEF (Methods). Fitting our data (Fig. 4) to equation (4) yields the limits of the $c'_{\mu\nu}$ parameters, which we report and compare with existing limits in Table 1. We improve the best measurements for those parameters, carried out by precision spectroscopy of dysprosium⁹, by up to two orders of magnitude, to a level of 10^{-18} . Alternatively, we can assume that Lorentz symmetry holds for electrons. We then can interpret our results as limits on Lorentz violation for photons (Methods) and improve on the bounds for Lorentz symmetry set by photon Michelson–Morley experiments² by up to five times (Table 1).

Our experimental scheme is readily applicable to other trapped-ion species considered for quantum information purposes. Many of those possess a long-lived electronic state with a non-vanishing angular momentum. Thus, further improvement can be achieved by increasing the Ramsey durations using metastable states with significantly longer lifetimes, such as 30 s for the barium ion²⁸, or by using ions more sensitive to Lorentz violation, such as highly charged ions²⁹. Additionally, by preparing a pure entangled state of the ions instead of a mixed state, it is possible to gain another factor of two in signal-to-noise ratio¹². Finally, we do not see any signature of limiting systematic effects, and thus expect that future extensions of our experimental technique with better statistics will yield tests of Lorentz symmetry at the level of 10^{-20} and below, where the polarization of black-body radiation in combination with temperature changes is expected to become relevant.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 August; accepted 12 November 2014.

- Michelson, A. A. & Morley, E. W. On the relative motion of the Earth and the luminiferous ether. *Am. J. Sci.* **34**, 333–345 (1887).
- Herrmann, S. *et al.* Rotating optical cavity experiment testing Lorentz invariance at the 10^{-17} level. *Phys. Rev. D* **80**, 105011 (2009).
- Eisele, Ch., Nevsky, A. Yu. & Schiller, S. Laboratory test of the isotropy of light propagation at the 10^{-17} level. *Phys. Rev. Lett.* **103**, 090401 (2009).
- Hughes, V. W., Robinson, H. G. & Beltran-Lopez, V. Upper limit for the anisotropy of inertial mass from nuclear resonance experiments. *Phys. Rev. Lett.* **4**, 342–344 (1960).
- Drever, R. W. P. A search for anisotropy of inertial mass using a free precession technique. *Phil. Mag.* **6**, 683–687 (1961).
- Smicklas, M., Brown, J. M., Cheuk, L. W., Smullin, S. J. & Romalis, M. V. New test of local Lorentz invariance using a ^{21}Ne -Rb-K comagnetometer. *Phys. Rev. Lett.* **107**, 171604 (2011).
- Allmendinger, F. *et al.* Upper limit for the anisotropy of inertial mass from nuclear resonance experiments. *Phys. Rev. Lett.* **112**, 110801 (2014).
- Peck, S. K. *et al.* Upper limit for the anisotropy of inertial mass from nuclear resonance experiments. *Phys. Rev. A* **86**, 012109 (2012).

- Hohensee, M. A. *et al.* Limits on violations of Lorentz symmetry and the Einstein equivalence principle using radio-frequency spectroscopy of atomic dysprosium. *Phys. Rev. Lett.* **111**, 050401 (2013).
- Altschul, B. Testing electron boost invariance with 2S–1S hydrogen spectroscopy. *Phys. Rev. D* **81**, 041701(R) (2010).
- Matveev, A. *et al.* Precision measurement of the hydrogen 1S–2S frequency via a 920-km fiber link. *Phys. Rev. Lett.* **110**, 230801 (2013).
- Roos, C. F., Chwalla, M., Kim, K., Riebe, M. & Blatt, R. 'Designer atoms' for quantum metrology. *Nature* **443**, 316–319 (2006).
- Kostecký, V. A. & Potting, R. CPT, strings, and meson factories. *Phys. Rev. D* **51**, 3923–3935 (1995).
- Kostecký, V. A. & Samuel, S. Spontaneous breaking of Lorentz symmetry in string theory. *Phys. Rev. D* **39**, 683–685 (1989).
- Hořava, P. Quantum gravity at a Lifshitz point. *Phys. Rev. D* **79**, 084008 (2009).
- Pospelov, M. & Shang, Y. Lorentz violation in Horava-Lifshitz-type theories. *Phys. Rev. D* **85**, 105001 (2012).
- Nibbelink, S. G. & Pospelov, M. Lorentz violation in supersymmetric field theories. *Phys. Rev. Lett.* **94**, 081601 (2005).
- Liberati, S. & Mattingly, D. Lorentz breaking effective field theory models for matter and gravity: theory and observational constraints. Preprint at <http://arxiv.org/abs/1208.1071> (2012).
- Colladay, D. & Kostecký, V. A. CPT violation and the standard model. *Phys. Rev. D* **55**, 6760–6774 (1997).
- Colladay, D. & Kostecký, V. A. Lorentz-violating extension of the standard model. *Phys. Rev. D* **58**, 116002 (1998).
- Russell, N. & Kostecký, V. A. Data tables for Lorentz and CPT violation. *Rev. Mod. Phys.* **83**, 11–31 (2011).
- Müller, H., Herrmann, S., Saenz, A., Peters, A. & Lämmerzahl, C. Optical cavity tests of Lorentz invariance for the electrons. *Phys. Rev. D* **68**, 116006 (2003).
- Müller, H. Testing Lorentz invariance by the use of vacuum and matter filled cavity resonators. *Phys. Rev. D* **71**, 045004 (2005).
- Kostecký, V. A. & Lane, C. Constraints on Lorentz violation from clock-comparison experiments. *Phys. Rev. D* **60**, 116010 (1999).
- Chwalla, M. *et al.* Precision spectroscopy with two correlated atoms. *Appl. Phys. B* **89**, 483–488 (2007).
- Chou, C. W., Hume, D. B., Rosenband, T. & Wineland, D. J. Optical clocks and relativity. *Science* **329**, 1630–1633 (2010).
- Madej, A., Dubé, P., Zhou, Z., Bernard, J. E. & Gertsulov, M. $^{88}\text{Sr}^+$ 445-THz single-ion reference at the 10^{-17} level via control and cancellation of systematic uncertainties and its measurement against the SI second. *Phys. Rev. Lett.* **109**, 203002 (2012).
- Iskrenova-Tchoukova, E. & Safronova, M. S. Theoretical study of lifetimes and polarizabilities in Ba^+ . *Phys. Rev. A* **78**, 012508 (2008).
- Safronova, M. S. *et al.* Highly-charged ions for atomic clocks, quantum information, and search for α -variation. *Phys. Rev. Lett.* **113**, 030801 (2014).

Acknowledgements This work was supported by the NSF CAREER programme grant no. PHY 0955650 and NSF grants no. PHY 1212442 and no. PHY 1404156, and was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. We thank H. Müller for critical reading of the manuscript.

Author Contributions H.H., M.A.H. and T.P. had the idea for the experiment. T.P. and M.R. carried out the measurements. S.G.P., I.I.T. and M.S.S. calculated the sensitivity of the energy to Lorentz violation. T.P., M.A.H. and H.H. wrote the main part of the manuscript. S.G.P., I.I.T. and M.S.S. wrote the Methods section on calculating the energy shift. All authors contributed to the discussions of the results and manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.P. (thaned.pruttivarasin@riken.jp) or H.H. (hhaeffner@berkeley.edu).

METHODS

Lorentz-violation parameters of electrons and photons. While Lorentz symmetry, or local Lorentz invariance, requires that the laws of physics be the same in all coordinate systems in the group formed by Lorentz transformations, it does not restrict our initial choice of coordinates. As a result, some forms of Lorentz violation cannot be unambiguously attributed to a single species of elementary particle without first specifying this coordinate choice. In particular, we can select our initial coordinates such that $c_{\mu\nu}$ (or its gauge field analogue $k_{\mu\nu}$) vanishes at leading order for any single species of particle (or gauge field). This particle then becomes a Lorentz-covariant ‘yardstick’ which other species can be compared against. For instance one might use light as the yardstick; that is, one would measure space such that $x_i = c_i t$ with the speed of light c_i constant in all three spatial directions i . Alternatively, one might use a coordinate system for which Lorentz symmetry is preserved for electrons. Then the value of $c_i t$ might not be the same in all three spatial directions. In this case Lorentz violation would manifest itself by breaking the rotational symmetry of the photon-mediated Coulomb force, yielding the same measurable energy shift as in the previous case. Consequently, a single experimental approach typically constrains a linear combination of particles and gauge fields.

To analyse which linear combination in the SME we test in our experiment, we neglect contributions of the nucleus to the Lorentz-violation signal for two reasons. First, the quadrupole moment of the doubly magic $^{40}\text{Ca}^+$ nucleus is expected to vanish. Second, the violations of Lorentz symmetry for nucleon constituents have been constrained to 10^{-26} for protons³⁰ and 10^{-29} for neutrons^{6,7}. In addition to the Lorentz-violating Lagrangian for electrons (equation (1)) in the SME, the Lorentz violation for electromagnetic fields (photons) is given by the parameters \tilde{k}^{JK} ($J, K = X, Y, Z$) (which are functions of $k_{\mu\nu}$) in the Lagrangian³¹

$$\mathcal{L} = \frac{1}{2} \left[(1 + \tilde{k}_{\text{tr}}) |\mathbf{E}|^2 - (1 - \tilde{k}_{\text{tr}}) |\mathbf{B}|^2 \right] + \frac{1}{2} [\mathbf{E} \tilde{\mathbf{k}}_{e-} \mathbf{E} - \mathbf{B} \tilde{\mathbf{k}}_{e-} \mathbf{B}] + \mathbf{E} \tilde{\mathbf{k}}_{o+} \mathbf{B} \quad (5)$$

where \tilde{k}_{tr} is a scalar, $\tilde{\mathbf{k}}_{e-}$ is a 3×3 traceless symmetric matrix that characterizes the anisotropy of the speed of light, and $\tilde{\mathbf{k}}_{o+}$ is a 3×3 antisymmetric matrix. By means of a coordinate transformation, observation of Lorentz violation for both the electrons and photons can be made to appear only in either the SME Lagrangian for the electron sector (equation (1)) or the photon sector (equation (5)). In both cases, the linear combinations of parameters relevant to our experiment are³²

$$\begin{aligned} c'_{X-Y} &= c_{XX} - c_{YY} + \frac{1}{2} (\tilde{k}_{e-}^{XX} - \tilde{k}_{e-}^{YY}) \\ c'_{XY} &= c_{XY} + \frac{1}{2} \tilde{k}_{e-}^{XY} \\ c'_{XZ} &= c_{XZ} + \frac{1}{2} \tilde{k}_{e-}^{XZ} \\ c'_{YZ} &= c_{YZ} + \frac{1}{2} \tilde{k}_{e-}^{YZ} \end{aligned}$$

The best existing limits on \tilde{k}_{e-}^{XY} , \tilde{k}_{e-}^{XZ} , \tilde{k}_{e-}^{YZ} and $\tilde{k}_{e-}^{XX} - \tilde{k}_{e-}^{YY}$ are given in ref. 2. In Table 1, we compare our results to these limits.

Experimental set-up. We trap a pair of $^{40}\text{Ca}^+$ ions in a linear Paul trap with an interelectrode distance of 1.0 mm. We apply a radio-frequency with peak-to-peak voltage of ~ 500 V to each pair of the radio-frequency electrodes. One pair of the electrodes is driven in antiphase with the other pair. With a ~ 4 V d.c. voltage applied across the end caps, we obtain trap frequencies of $2\pi \times (2.2, 2.0)$ MHz in the radial directions and $2\pi \times 210$ kHz in the axial direction. The axial direction is aligned horizontally in the laboratory frame. To define a quantization axis, we apply a static magnetic field of 3.930 G vertically (at 45° with respect to both radial directions of the trap) using a coil. Additionally, we use another magnetic coil to compensate for residual magnetic field gradient along the axial direction.

Two independent 729 nm laser beams in the vertical direction drive π - and $\pi/2$ -pulses on the $S_{1/2}$ - $D_{5/2}$ transition on each ion separately. Both beams are derived from a laser stabilized to a high-finesse optical cavity to better than 100 Hz. Another beam path addressing both ions in the horizontal direction (at 45° with respect to the axial direction) is used for Doppler cooling (397 nm and 866 nm) and repumping the $D_{5/2}$ state (854 nm). We perform all laser light switching and frequency shifting using acousto-optical modulators (AOMs) in a double-pass configuration. We generate all radio-frequency voltages supplied to the AOMs using direct-digital-synthesizer chips from Analogue Devices (AD9910). The timing in the experimental sequence is controlled by a field-programmable gate-array (FPGA) module (XEM6010) from Opal Kelly. We characterize the stability of the on-board crystal oscillator using a frequency counter (Agilent 53210A). The clock stability is measured to be at the level of 4×10^{-7} , which translates to better than 5 μHz stability in the oscillation signal of the measurement of Lorentz violation.

Measurement scheme. The experimental sequence is shown in Fig. 3. We measure four independent oscillation signals for the two states $|\Psi^L\rangle \equiv (1/\sqrt{2})(|+5/2, -5/2\rangle + |+1/2, -1/2\rangle)$ and $|\Psi^R\rangle \equiv (1/\sqrt{2})(|-5/2, +5/2\rangle + |-1/2, +1/2\rangle)$, each with both short ($T_{\text{short}} = 5$ ms) and long ($T_{\text{long}} = 100$ ms) Ramsey duration (Fig. 3b). Within each measurement block in Fig. 3b, the order in which we perform Ramsey spectroscopy for each state and the Ramsey duration is randomized to average out systematic noise that might coincide with the period (~ 60 s) of the measurement block.

In general, the oscillation signal has the form $S(t) = \mathcal{A} \cos(2\pi f t + \phi_{\text{offset}} + \phi_{\text{laser}}) + \mathcal{B}$, where \mathcal{A} is the amplitude of the signal, \mathcal{B} is a possible offset to the overall level of the signal, f is the oscillation frequency, ϕ_{offset} is the phase offset and ϕ_{laser} is an additional phase that we can control by changing the phase of the 729 nm laser (using the radio-frequency signal supplied to the AOM for each beam path) that drives π - and $\pi/2$ -pulses on the $S_{1/2}$ - $D_{5/2}$ transition of the ions.

For a given state and Ramsey duration, the Ramsey interferometric cycle shown in Fig. 3a is repeated 200 times. To cancel out drifts in the offset of the signal, \mathcal{B} , we perform the first 100 cycles of the Ramsey sequence with the phase of the laser given by ϕ_{laser} and the next 100 cycles with the phase of the laser given by $\phi_{\text{laser}} + \pi$. We then calculate the difference between these two signals, $(S(\phi_{\text{laser}}) - S(\phi_{\text{laser}} + \pi))/2 = \mathcal{A} \cos(2\pi f t + \phi_{\text{offset}} + \phi_{\text{laser}})$, which does not depend on \mathcal{B} .

For a fixed Ramsey duration T , the oscillation signal $S(T) = \mathcal{A} \cos(2\pi f T + \phi_{\text{offset}} + \phi_{\text{laser}})$ is most sensitive to variation in the oscillation frequency, f , when the signal crosses zero, that is, when $2\pi f T + \phi_{\text{offset}} + \phi_{\text{laser}} = \pi/2$. We make sure that the oscillation signal remains close to zero by adding the phase correction calculated from the oscillation signal, that is, $\delta\phi = \cos^{-1}(S(T)/\mathcal{A}) - \pi/2$, to the phase of the laser light, ϕ_{laser} . The long-term measurement of the variation in the oscillation frequency, δf , is then derived from the phase correction data using $\delta\phi = 2\pi T \delta f$.

In addition to the change in the oscillation frequency, any change in ϕ_{offset} in the state preparation affects the phase correction: $\delta\phi = 2\pi T \delta f + \delta\phi_{\text{offset}}$. To correct for a contribution from this phase offset, we use signals from the two Ramsey durations ($T_{\text{short}} = 5$ ms and $T_{\text{long}} = 100$ ms) and calculate the difference between the phase corrections: $\delta\phi_{\text{long}} - \delta\phi_{\text{short}} = 2\pi(T_{\text{long}} - T_{\text{short}})\delta f$. The oscillation frequency for the state $|\Psi^{L,R}\rangle$ is given by $\delta f_{L,R} = [(\delta\phi_{\text{long}} - \delta\phi_{\text{short}})/2\pi(T_{\text{long}} - T_{\text{short}})]_{L,R}$, where the effective Ramsey duration is $T_{\text{long}} - T_{\text{short}} = 95$ ms.

Whereas the linear Zeeman effects from a magnetic field common to both ions drops out, the linear Zeeman effect due to a magnetic field gradient does not. In a typical unshielded laboratory environment, the gradient remains stable enough to allow for contrast with Ramsey times of about 30 s (ref. 33). To remove extant frequency variations from the gradient, we take the average frequency $\delta\bar{f} = (\delta f_L + \delta f_R)/2$ (Extended Data Fig. 1), which now contains only contributions from the electric quadrupole shift, the quadratic Zeeman shift, a.c. Stark shifts from oscillating trapping fields, and shifts from Lorentz violation.

We characterize the effect of the electric quadrupole shift by measuring the oscillation frequency $\delta\bar{f}$ as a function of the electric field gradient by changing the axial trap frequency. For our experimental set-up, we obtain $\delta\bar{f} = 4.0(8)$ (Hz mm² V⁻¹) $E' + 8.9(8)$ (Hz), where E' is the electric field gradient. At our operating axial trap frequency of 210 kHz, this translates to variations in the quadrupole shift due to changes in the axial trap frequency of 27 ± 12 mHz kHz⁻¹. The offset of 8.9(8) Hz is due to the quadratic Zeeman shift, which agrees with the estimated value of 8 Hz for the applied magnetic field of 3.930 G. Any change in the magnitude of the applied magnetic field near our operating value of 3.930 G gives a variation of the quadratic Zeeman shift of 4 mHz mG⁻¹. Using the ions as a probe, we measure both the magnetic field and the axial trap frequency during the course of the experiment and correct for their contributions to the oscillation signal. Over the course of our 23 h-long run, our axial trap frequency varies by less than ~ 1 kHz and the magnetic field by less than 1 mG. These instabilities translate into variations in the correction to the oscillation frequency of ~ 30 mHz, due to the quadrupole shift, and 3 mHz, due to the magnetic field. Fitting the model in equation (4) to the corrections only, we find that not taking into account the axial frequency instability would cause a false Lorentz-violation signal with amplitudes of less than 3 mHz, and that not correcting for the magnetic field instabilities would cause a signal with amplitudes of less than 0.5 mHz. Thus, in principle no correction for their drift would have been necessary. We note also that by measuring those quantities during the measurement run, their average contributions are expected to decrease as fast as the primary measurement signal, and should thus pose no limitation on improved Lorentz symmetry tests with longer measurement runs.

The oscillating electric field from the radio-frequency electrodes of the trap induces a.c. Stark shifts of the atomic transitions of the ions. The amplitude of the oscillating field experienced by the ions depends on the stray background static electric field. For our trap, we estimate that the stray electric field at the vicinity of the ions is ~ 5 V cm⁻¹. This produces a differential a.c. Stark shift between the $|\pm 1/2\rangle$ and $|\pm 5/2\rangle$ states to be ~ 120 mHz (ref. 34). The stray field is expected to be stable

at better than the 10^{-2} level during the course of the experiment, which translates to a change of less than 4 mHz in the oscillation frequency for the two-ion state. **Statistical analysis of the data.** After each measurement block (Fig. 3b), we obtain a data point for the frequency difference between both states. We then bin the data points within 60 min intervals. The error bar for each binned data point is assigned using the calculated standard deviation within each bin. To extract the amplitudes of Lorentz violation, we perform a weighted least-squares fit of the binned data points to the model given in equation (4). We scale the 1 s.d. errors of the fitted parameters with $\sqrt{\chi^2_{\text{reduced}}} = 1.3$ to account conservatively for other remaining systematics.

Calculation of the energy shift due to the Lorentz-violation for $^{40}\text{Ca}^+$. Violations of Lorentz symmetry and Einstein's equivalence principle in bound electronic states result in a small shift of the Hamiltonian that can be described by⁹

$$\delta\mathcal{H} = -\left(C_0^{(0)} - \frac{2U}{3c^2}c'_{00}\right)\frac{\mathbf{p}^2}{2} - \frac{1}{6}C_0^{(2)}T_0^{(2)}$$

where we use atomic units, \mathbf{p} is the momentum of a bound electron, U is the Newtonian potential and c is the speed of light. The parameters $C_0^{(0)}$, $C_0^{(2)}$ and c'_{00} are elements of the $c'_{\mu\nu}$ tensor, which characterizes Lorentz violation. The relativistic form of the \mathbf{p}^2 operator is $c_{\gamma 0} \gamma^j p_j$ (a summation is implied by repeated indices), where γ^j are the Dirac gamma matrices. The non-relativistic form of the $T_0^{(2)}$ operator is $T_0^{(2)} = \mathbf{p}^2 - 3p_z^2$, where p_z is the component of the momentum along the quantization axis, and the relativistic form is $T_0^{(2)} = c_{\gamma 0} (\gamma^j p_j - 3\gamma^3 p_3)$. Therefore, the shift in the $\text{Ca}^+ 3d^2D_{5/2}$ energy level due to the $c'_{\mu\nu}$ tensor depends on the values of the $\langle 3d^2D_{5/2} | \mathbf{p}^2 | 3d^2D_{5/2} \rangle$ and $\langle 3d^2D_{5/2} | T_0^{(2)} | 3d^2D_{5/2} \rangle$ matrix elements.

Using the Wigner-Eckart theorem we express the matrix element of the irreducible tensor operator $T_0^{(2)}$ through the reduced matrix element $\langle J || T^{(2)} || J \rangle$ of the operator $T^{(2)}$ as

$$\langle J m_J | T_0^{(2)} | J m_J \rangle = \frac{-J(J+1) + 3m_J^2}{\sqrt{(2J+3)(J+1)(2J+1)(2J-1)}} \langle J || T^{(2)} || J \rangle \quad (6)$$

The expressions for the \mathbf{p}^2 and $T^{(2)}$ matrix elements are given in the supplementary material of ref. 9. The values of the angular factor in equation (6) (that is, the prefactor of the reduced matrix element) are $-0.27951 + 0.22361 m_J^2$ for $3d^2D_{3/2}$ and $-0.21348 + 0.073193 m_J^2$ for $3d^2D_{5/2}$.

First we calculated the required matrix elements in a lowest-order Dirac-Fock (DF) approximation and then with an additional random-phase approximation (DF+RPA). Next we carried out much more accurate calculations using the configuration interaction method with single and double excitations (CI+SD) and four variants of the all-order (linearized coupled-cluster) method³⁵. The virial theorem is also used for the \mathbf{p}^2 calculations.

The results are summarized in Extended Data Table 1. We note that we list the reduced matrix elements for the $T^{(2)}$ operator but actual matrix elements for the \mathbf{p}^2 operator because there is no necessity to introduce reduced matrix elements for a scalar operator. The values in columns DF(FC) and DF are lowest-order Dirac-Fock values calculated with and without the frozen-core approximation. In the frozen-core approximation the Dirac-Fock equations for the core electrons are solved self-consistently first and the valence orbital is calculated with an unchanged, that is, 'frozen', core. For the \mathbf{p}^2 operator such an approximation appears to give very poor results for the $3d$ states. If the core orbitals are allowed to vary together with the valence orbital, the lowest-order value differs by only 16% from the final virial theorem value. Addition of the RPA correction to the frozen-core Dirac-Fock value fixes this problem as well, because RPA corrections describe the reaction of the core electrons to an externally applied perturbation. The perturbation produced by the operator \mathbf{p}^2 is very large and, as a result, the RPA corrections for $\langle \psi | \mathbf{p}^2 | \psi \rangle$ matrix elements are large. Such a problem does not arise for the $T^{(2)}$ operator; the correlation correction to its matrix elements is much smaller and the accuracy of the resulting values is much higher.

The CI+SD calculations are carried out using the Dirac-Fock basis for the occupied core and valence atomic states and the Dirac-Fock-Sturm basis for unoccupied virtual orbitals; the frozen-core approximation is not used. The description of the Dirac-Fock-Sturm equations is given in refs 36, 37. The configuration state functions are constructed from the one-electron wavefunctions as a linear combination of Slater determinants. The set of the configuration state functions is generated including all single and double excitations into one-electron states of the positive spectrum. Single excitations are allowed to all core shells; double excitations are allowed to $3s$ and $3p$ core shells.

To calculate the value $\langle \psi | \mathbf{p}^2 | \psi \rangle$, where $|\psi\rangle$ is the valence electron wavefunction, we also used the approach based on the virial theorem. In the nonrelativistic limit the virial theorem can be written in the form

$$E = -\frac{1}{2} \langle \Psi | \sum_{i=1}^N \mathbf{p}^2(i) | \Psi \rangle$$

where E is the total energy of the system, N is the total number of electrons and $|\Psi\rangle$ is the total wavefunction of all electrons in the atom. Therefore, the value $\langle \psi | \mathbf{p}^2 | \psi \rangle$ can be calculated using the removal energies of the valence electron. The virial theorem makes it possible to calculate the expectation value of the \mathbf{p}^2 operator as twice the difference of the total energies E_N and E_{N-1} of N and $N-1$ systems. Because the differential energy E can be calculated with an accuracy much higher than the wavefunction Ψ , this approach is appropriate for the light atoms and ions where relativistic effects are negligible. The virial theorem results that use experimental data for the $3d$ removal energies from ref. 38 are listed in column VT in Extended Data Table 1.

We have also carried out the calculations of the $\langle \psi | \mathbf{p}^2 | \psi \rangle$ and $\langle \psi | T^{(2)} | \psi \rangle$ matrix elements using the all-order (linearized coupled-cluster) method³⁵. The all-order method gave very accurate values of the $3d_j$ lifetimes³⁹ and quadrupole moments⁴⁰ in a Ca^+ ion. In the all-order method, single, double and partial triple excitations of Dirac-Hartree-Fock wavefunctions are included to all orders of perturbation theory. We refer the reader to the review in ref. 35 for the description of the all-order method and its applications. Both single-double (SD) and single-double-partial-triple (SDpT) *ab initio* all-order calculations were carried out. In addition, a scaling of the dominant terms³⁵ was carried out for both SD and SDpT calculations to improve the accuracy and to evaluate the uncertainty of the final values. The calculations were carried out with both nonrelativistic and relativistic operators; the differences were found to be negligible at the present level of accuracy. The values calculated with relativistic operators are listed in Extended Data Table 1.

The virial theorem values are taken as final for the matrix element of the \mathbf{p}^2 operator. The uncertainty of 12% is estimated as the difference of the virial theorem and all-order values. The SD scaled values are taken as final for the $T^{(2)}$ operator (see refs 39, 40 for the discussion of the choice of the final all-order values). The uncertainty is determined as the spread of the four all-order values. On substituting the final all-order values of the $\langle 3d^2D_{5/2} | c_{\gamma 0} (\gamma^j p_j - 3\gamma^3 p_3) | 3d^2D_{5/2} \rangle$ matrix element into equation (6) and using the virial theorem value of $\langle 3d^2D_{5/2} | \mathbf{p}^2 | 3d^2D_{5/2} \rangle$, we get

$$\frac{\Delta E}{h} \approx \left(C_0^{(0)} - \frac{2U}{3c^2}c'_{00} \right) \times (-2.46 \times 10^{15} \text{ Hz}) + C_0^{(2)} \times (2.17 \times 10^{15} - 1.47 \times 10^{15} m_J^2) \text{ Hz}$$

for $3d^2D_{3/2}$, and

$$\frac{\Delta E}{h} \approx \left(C_0^{(0)} - \frac{2U}{3c^2}c'_{00} \right) \times (-2.46 \times 10^{15} \text{ Hz}) + C_0^{(2)} \times (2.16 \times 10^{15} - 7.42 \times 10^{14} m_J^2) \text{ Hz}$$

for $3d^2D_{5/2}$, where the uncertainties in the frequency coefficients of $\left(C_0^{(0)} - \frac{2U}{3c^2}c'_{00} \right)$ and $C_0^{(2)}$ are estimated to be 12% and 2%, respectively, and the atomic units are converted to SI units using $1 \text{ a.u.} \approx h \times (6.57968 \times 10^{15} \text{ Hz})$.

The frequency difference (in Hz) between the shifts of the $m_J = 5/2$ and $m_J = 1/2$ states for a pair of $^{40}\text{Ca}^+$ ions used in our experiment is given by

$$\frac{2}{h} (E_{m_J=5/2} - E_{m_J=1/2}) = (-1.484 \times 10^{15} \text{ Hz}) \times ((5/2)^2 - (1/2)^2) C_0^{(2)} = (-8.9(2) \times 10^{15} \text{ Hz}) \times C_0^{(2)}$$

Frame dependence of the $c'_{\mu\nu}$ tensor. Because of the Earth's motion, $c'_{\mu\nu}$ in the local laboratory frame varies according to the time-dependent Lorentz transformation given by

$$c'_{\mu\nu} = c_{MN} A_{\mu}^M A_{\nu}^N$$

where A is the Lorentz transformation matrix and c'_{MN} is $c'_{\mu\nu}$ written in the Sun-centred, celestial-equatorial frame (SCCEF). The matrix A consists of a rotation and a velocity boost of the experiment with respect to the Sun. In the laboratory frame, we define the \hat{x} axis to point to the East, the \hat{y} axis to point to the North and the \hat{z} axis to point upward. The rotation matrix that transforms from the SCCEF to the local laboratory frame is given by

$$R = \begin{pmatrix} -\sin(\omega_{\oplus} T) & \cos(\omega_{\oplus} T) & 0 \\ -\cos(\chi) \cos(\omega_{\oplus} T) & -\cos(\chi) \sin(\omega_{\oplus} T) & \sin(\chi) \\ \sin(\chi) \cos(\omega_{\oplus} T) & \sin(\chi) \sin(\omega_{\oplus} T) & \cos(\chi) \end{pmatrix}$$

where the angle $\chi \approx 52.1^\circ$ is the colatitude of the experiment (Berkeley, California), T is time since the vernal equinox of 2014 and $\omega_{\oplus} = 2\pi/(23.93 \text{ h})$ is the sidereal

angular frequency of the Earth's rotation. The boost of the experiment in the SCCEF is given by

$$\beta = \begin{pmatrix} -\beta_{\oplus} \sin(\eta) \cos(\Omega T) \\ \beta_{\oplus} \cos(\eta) \cos(\Omega T) - \beta_L \sin(\chi) \cos(\omega_{\oplus} T) \\ -\beta_{\oplus} \sin(\Omega T) + \beta_L \sin(\chi) \sin(\omega_{\oplus} T) \end{pmatrix}$$

where $\beta_{\oplus} \approx 10^{-4}$ is the boost from the Earth's orbital velocity, $\beta_L \approx 1.5 \times 10^{-6}$ is the boost from the Earth's rotation, Ω is the yearly sidereal angular frequency and $\eta \approx 23.4^\circ$ is the angle between the ecliptic plane and the Earth's equatorial plane.

The parameter relevant to our experiment is $C_0^{(2)}$. With the Lorentz transformation applied to $c'_{\mu\nu}$ in the SCCEF, we can write the value of $C_0^{(2)}$ in the local laboratory frame in terms of $c'_{\mu\nu}$ in the SCCEF as

$$C_0^{(2)} = A + \sum_j (C_j \cos(\omega_j T) + S_j \sin(\omega_j T))$$

where the index j runs over all angular frequencies (ω_j) and the corresponding amplitudes (C_j, S_j) given in Extended Data Table 2, and A is a constant offset. For our 23 h measurement, the time-dependent Lorentz-violation signal is given at leading order by

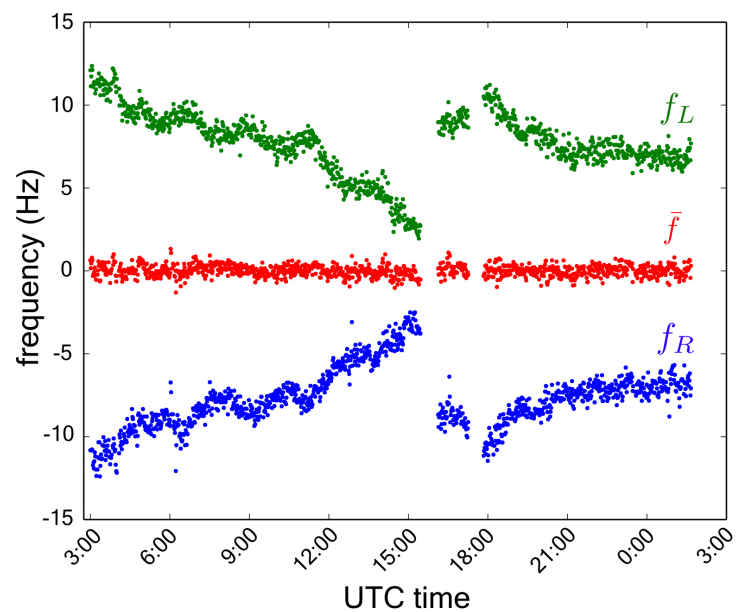
$$C_0^{(2)} = -3 \sin(2\chi) c'_{XZ} \cos(\omega_{\oplus} T) - 3 \sin(2\chi) c'_{YZ} \sin(\omega_{\oplus} T) \\ - \frac{3}{2} (c'_{XX} - c'_{YY}) \sin^2(\chi) \cos(2\omega_{\oplus} T) - 3 c'_{XY} \sin^2(\chi) \sin(2\omega_{\oplus} T)$$

We fit our binned 23 h measurement data to this model and extract Lorentz-violation parameters. In Table 1 we report uncorrelated combinations of parameters by diagonalizing the covariance matrix from the fit. We scale the 1 s.d. uncertainties

from the fit with $\sqrt{\chi^2_{\text{reduced}}} = 1.3$ to account conservatively for other remaining systematics.

With a year-long measurement, we expect to reach a sensitivity of 1 mHz in the ions' oscillation frequency. This level of sensitivity allows us to bound c'_{TX} , c'_{TY} and c'_{TZ} at the 10^{-16} level, which will improve the present limits^{9,41} on these parameters by at least an order of magnitude.

30. Wolf, P., Chapelet, F., Bize, S. & Clairon, A. Cold atom clock test of Lorentz invariance in the matter sector. *Phys. Rev. Lett.* **96**, 060801 (2006).
31. Kostelecký, V. A. & Mewes, M. Signals for Lorentz violation in electrodynamics. *Phys. Rev. D* **66**, 056005 (2002).
32. Bailey, Q. & Kostelecký, V. A. Lorentz-violating electrostatics and magnetostatics. *Phys. Rev. D* **70**, 076006 (2004).
33. Häffner, H. et al. Robust entanglement. *Appl. Phys. B* **81**, 151–153 (2005).
34. Yu, N., Zhao, X., Dehmelt, H. & Nagourney, W. Stark shift of a single barium ion and potential application to zero-point confinement in a rf trap. *Phys. Rev. A* **50**, 2738–2741 (1994).
35. Safronova, M. S. & Johnson, W. R. All-order methods for relativistic atomic structure calculations. *Adv. At. Mol. Opt. Phys.* **55**, 191–233 (2008).
36. Tupitsyn, I. I. et al. Magnetic-dipole transition probabilities in B-like and Be-like ions. *Phys. Rev. A* **72**, 062503 (2005).
37. Tupitsyn, I. I. et al. Relativistic calculations of the charge-transfer probabilities and cross sections for low-energy collisions of H-like ions with bare nuclei. *Phys. Rev. A* **82**, 042701 (2010).
38. Kramida, A. Ralchenko, Yu., Reader, J. & NIST ASD Team. *NIST Atomic Spectra Database (version 5.1)* <http://physics.nist.gov/asd> (NIST, 2013).
39. Kreuter, A. et al. Experimental and theoretical study of the $3d^2D$ -level lifetimes of $^{40}\text{Ca}^+$. *Phys. Rev. A* **71**, 032504 (2005).
40. Jiang, D., Arora, B. & Safronova, M. S. Electric quadrupole moments of metastable states of Ca^+ , Sr^+ , and Ba^+ . *Phys. Rev. A* **78**, 022514 (2008).
41. Altschul, B. Limits on Lorentz Violation from synchrotron and inverse Compton sources. *Phys. Rev. Lett.* **96**, 201101 (2006).



Extended Data Figure 1 | Cancellation of the contributions from the magnetic field gradient. The frequency measurements of the states $|\psi^L\rangle$ and $|\psi^R\rangle$ for a Ramsey duration of 100 ms are shown in the top green (f_L) and

bottom blue (f_R) data sets, respectively. We offset both data sets for visualization purposes. The contribution from the magnetic field gradient is subtracted out in the average frequency $\bar{f} = (f_L + f_R)/2$, which is shown as red data points.

Extended Data Table 1 | Lowest-order DF, DF+RPA, CI+SD and all-order results for the $\langle 3d^2D_J | p^2 | 3d^2D_J \rangle$ and $\langle 3d^2D_J | T^{(2)} | 3d^2D_J \rangle$ matrix elements in Ca^+ in atomic units

Matrix element	DF(FC)	DF	RPA	CI+SD	All-order	VT	Final
$\langle 3d^2D_{3/2} p^2 3d^2D_{3/2} \rangle$	3.05	0.67	0.66	0.73	0.83	0.748	0.75(9)
$\langle 3d^2D_{5/2} p^2 3d^2D_{5/2} \rangle$	3.04	0.66	0.66	0.73	0.83	0.748	0.75(9)
$\langle 3d^2D_{3/2} T^{(2)} 3d^2D_{3/2} \rangle$	5.45	6.22	5.72	6.89	7.09		7.09(12)
$\langle 3d^2D_{5/2} T^{(2)} 3d^2D_{5/2} \rangle$	7.12	8.11	7.47	8.98	9.25		9.25(15)

The virial theorem values are listed in column VT. The values in columns DF(FC) and DF are lowest-order Dirac–Fock values calculated with and without the frozen-core approximation.

Extended Data Table 2 | Amplitudes of various frequency components for $C_0^{(2)}$ expressed in terms of $c'_{\mu\nu}$ in the SCCEF

ω_j	C_j	S_j
ω_{\oplus}	$-3 \sin(2\chi)c'_{XZ} + 2c'_{TY}\beta_L$	$-3 \sin(2\chi)c'_{YZ} - 2c'_{TX}\beta_L$
$2\omega_{\oplus}$	$-\frac{3}{2}(c'_{XX} - c'_{YY})\sin^2(\chi)$	$-3c'_{XY}\sin^2(\chi)$
Ω	$-\frac{1}{2}\beta_{\oplus}(3 \cos(2\chi) + 1)(c'_{TY} \cos(\eta) - 2c'_{TZ} \sin(\eta))$	$\frac{1}{2}\beta_{\oplus}c'_{TX}(3 \cos(2\chi) + 1)$
2Ω	0	0
$\Omega - \omega_{\oplus}$	$\frac{3}{2}\beta_{\oplus}c'_{TX} \sin(\eta) \sin(2\chi)$	$-\frac{3}{2}\beta_{\oplus} \sin(2\chi) (c'_{TY} \sin(\eta) + c'_{TZ}(1 + \cos(\eta)))$
$\Omega + \omega_{\oplus}$	$\frac{3}{2}\beta_{\oplus}c'_{TX} \sin(\eta) \sin(2\chi)$	$-\frac{3}{2}\beta_{\oplus} \sin(2\chi) (c'_{TZ}(1 - \cos(\eta)) - c'_{TY} \sin(\eta))$
$2\Omega - \omega_{\oplus}$	0	0
$2\Omega + \omega_{\oplus}$	0	0
$\Omega - 2\omega_{\oplus}$	$-3\beta_{\oplus}c'_{TY} \cos^2\left(\frac{\eta}{2}\right) \sin^2(\chi)$	$-3\beta_{\oplus}c'_{TX} \cos^2\left(\frac{\eta}{2}\right) \sin^2(\chi)$
$\Omega + 2\omega_{\oplus}$	$3\beta_{\oplus}c'_{TY} \sin^2\left(\frac{\eta}{2}\right) \sin^2(\chi)$	$-3\beta_{\oplus}c'_{TX} \sin^2\left(\frac{\eta}{2}\right) \sin^2(\chi)$
$2\Omega - 2\omega_{\oplus}$	0	0
$2\Omega + 2\omega_{\oplus}$	0	0

The frequencies ω_{\oplus} and Ω are the daily and yearly sidereal angular frequencies, respectively. The angle $\chi \approx 52.1^\circ$ is the colatitude of the experiment (Berkeley, California). The angle $\eta \approx 23.4^\circ$ is the angle between the plane of the ecliptic and the Earth's equatorial plane. $\beta_{\oplus} \approx 10^{-4}$ is the boost from the Earth's orbital velocity and $\beta_L \approx 1.5 \times 10^{-6}$ is the boost from the Earth's rotation. For our 23 h-measurement, contributions from these two boosts are negligible.

Anomalous dispersions of ‘hedgehog’ particles

Joong Hwan Bahng¹, Bongjun Yeom², Yichun Wang¹, Siu On Tung³, J. Damon Hoff⁴ & Nicholas Kotov^{1,2,3,5,6}

Hydrophobic particles in water and hydrophilic particles in oil aggregate, but can form colloidal dispersions if their surfaces are chemically camouflaged with surfactants, organic tethers, adsorbed polymers or other particles that impart affinity for the solvent and increase interparticle repulsion^{1,2}. A different strategy for modulating the interaction between a solid and a liquid uses surface corrugation, which gives rise to unique wetting behaviour^{3–5}. Here we show that this topographical effect can also be used to disperse particles in a wide range of solvents without recourse to chemicals to camouflage the particles’ surfaces: we produce micrometre-sized particles that are coated with stiff, nanoscale spikes and exhibit long-term colloidal stability in both hydrophilic and hydrophobic media. We find that these ‘hedgehog’ particles do not interpenetrate each other with their spikes, which markedly decreases the contact area between the particles and, therefore, the attractive forces between them. The trapping of air in aqueous dispersions, solvent autoionization at highly developed interfaces, and long-range electrostatic repulsion in organic media also contribute to the colloidal stability of our particles. The unusual dispersion behaviour of our hedgehog particles, overturning the notion that

like dissolves like, might help to mitigate adverse environmental effects of the use of surfactants and volatile organic solvents, and deepens our understanding of interparticle interactions and nanoscale colloidal chemistry.

We imparted strong corrugation onto the surface of carboxylated polystyrene microspheres (μ PSs) by attaching rigid zinc oxide (ZnO) nanoscale spikes (‘nanospikes’). This involves initial absorption of positively charged ZnO nanoparticles (NPs) onto the negatively charged μ PSs and subsequent growth of ZnO nanowires using established protocols⁶. The resultant hedgehog particles combine micro- and nanoscale structural features (Fig. 1a), and their geometrical and topographical specifications can be adjusted by changing the growth conditions to modify the surface densities, lengths and diameters of nanospikes and the μ PS diameters (Fig. 1b–e, Supplementary Information section 1 and Supplementary Figs 2–5).

As-made hedgehog particles, with their polar ZnO surfaces, are highly hydrophilic. They form excellent dispersions in water (Fig. 1f, l) and other hydrophilic solvents. When rendering the hedgehog particles hydrophobic by silanization of the ZnO nanospikes with (7-octen-1-yl)

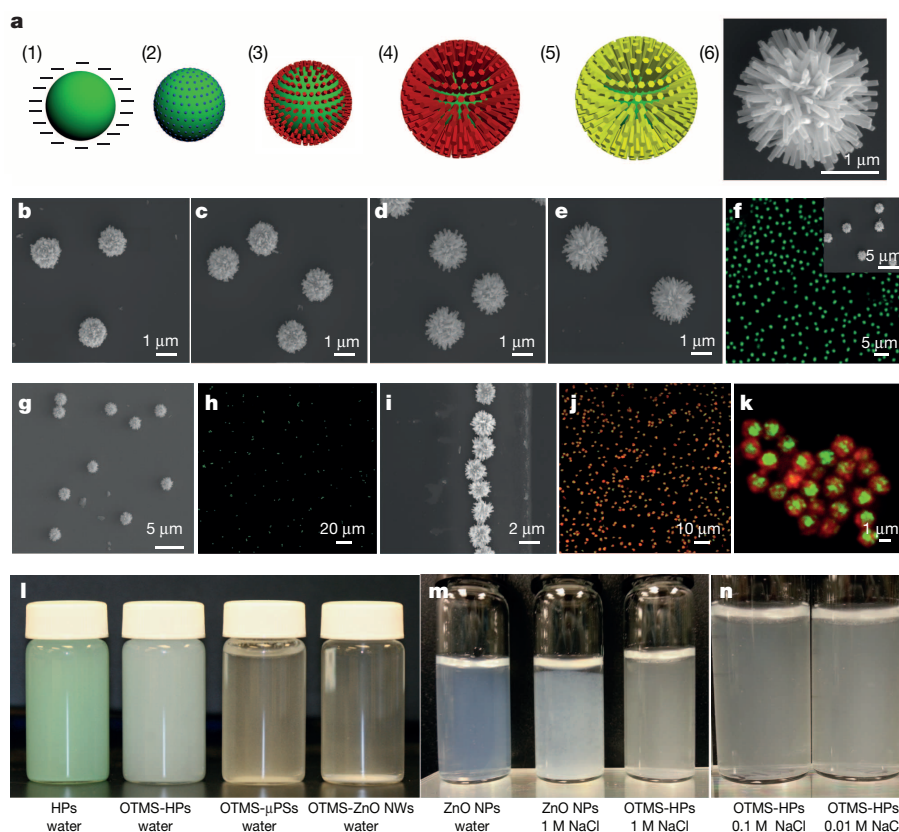


Figure 1 | Hedgehog particles. **a**, Negatively charged, carboxylate-terminated μ PSs are used as core templates (1) on which positively charged ZnO NPs are adsorbed (2). ZnO nanospikes are grown from ZnO nanoparticles (3) to a designed length (4, 6). Hedgehog particles are rendered hydrophobic by exposure to OTMS or PFTS (5). **b–e**, SEM images of hedgehog particles with different ZnO nanospike lengths: 0.19 μ m (b), 0.27 μ m (c), 0.4 μ m (d), 0.6 μ m (e). **f**, Confocal microscopy of an aqueous dispersion of hydrophilic hedgehog particles with fluorescently labelled μ PSs. Inset, SEM image for the same hedgehog particles. **g, h**, SEM (g) and confocal microscopy (h) of an aqueous dispersion of OTMS-HPs. **i**, SEM image of particles from the bulk of an aqueous OTMS-HP dispersion collected five days after initial preparation. **j, k**, Confocal microscopy images of fluorescent OTMS-HPs (green, $\lambda_{\text{max}} = 486$ nm) with adsorbed hydrophobic CdSe nanoparticles (red, $\lambda_{\text{max}} = 655$ nm) in an aqueous dispersion (j) and in the dried state (k). **l**, Photographs of aqueous dispersions of (left to right) hydrophilic hedgehog particles (HPs) with green-dyed μ PSs, OTMS-HPs, OTMS- μ PSs and OTMS-ZnO nanowires (NWs). **m**, Photographs of (left to right) ZnO nanoparticles (NPs) in water, ZnO nanoparticles in 1 M NaCl, and OTMS-HPs in 1 M NaCl. **n**, Photographs of OTMS-HPs in (left to right) 0.1 M NaCl and 0.01 M NaCl.

¹Department of Biomedical Engineering, University of Michigan, 1107 Carl A. Gerstacker Building, 2200 Bonisteel Boulevard, Ann Arbor, Michigan 48109, USA. ²Department of Chemical Engineering, University of Michigan, 3074 H.H. Dow Building, 2300 Hayward Street, Ann Arbor, Michigan 48109, USA. ³Macromolecular Science and Engineering Program, University of Michigan, 3062C H.H. Dow Building, 2300 Hayward Street, Ann Arbor, Michigan 48109, USA. ⁴Single Molecule Analysis in Real Time (SMART) Center, Ann Arbor, Michigan 48109, USA. ⁵Department of Material Science and Engineering, University of Michigan, 3074 H.H. Dow Building, 2300 Hayward Street, Ann Arbor, Michigan 48109, USA. ⁶Biointerfaces Institute, University of Michigan, North Campus Research Complex, 2800 Plymouth Road, Ann Arbor, Michigan 48109, USA.

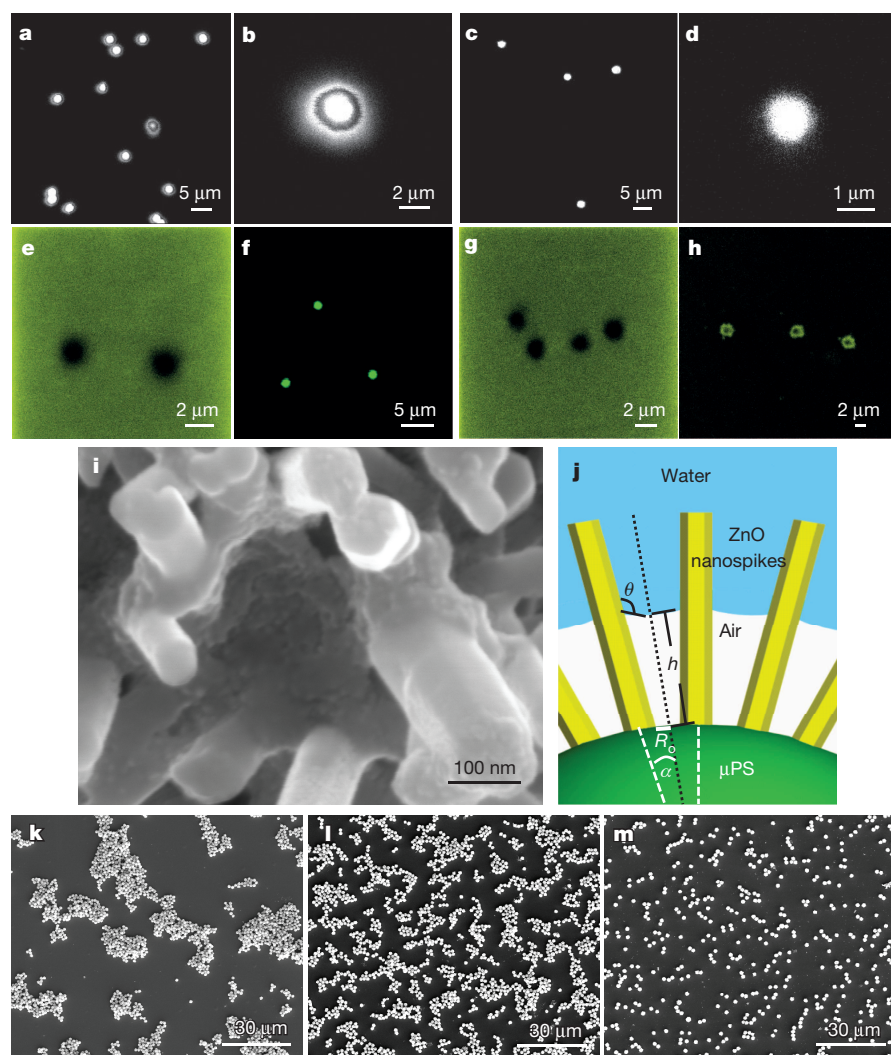


Figure 2 | Interface between hydrophobic hedgehog particles and water. **a–h**, Confocal microscopy images of hydrophobic hedgehog particles labelled with hydrophobic CdSe NPs in aqueous dispersions (**a**, **b**) and hydrophilic hedgehog particles labelled with hydrophilic CdTe NPs in aqueous dispersions (**c**, **d**); hydrophobic hedgehog particles in an aqueous solution containing hydrophilic TGA-stabilized CdTe nanoparticles with green ($\lambda_{\text{max}} = 540 \text{ nm}$) emission (**e**); hydrophilic hedgehog particles in an aqueous solution containing hydrophilic TGA-stabilized CdTe nanoparticles (**f**); the same sample from image **e** after five days of storage in dark (**g**); and the same sample from image **f** after five days of storage in dark (**h**). **i**, SEM image of a hydrophobic hedgehog particle with a self-assembled film of TGA-depleted CdTe nanoparticles between the ZnO nanospikes, indicating the location of the air–water interface. The hydrophobic hedgehog particles were immersed in an aqueous solution of CdTe nanoparticles for 72 h. **j**, Schematic diagram of the air–water interface, showing the experimental parameters (definitions in Supplementary Information). **k–m**, SEM images of aqueous dispersions of hydrophobic hedgehog particles with ZnO nanospike lengths of 0.19 μm (**k**), 0.40 μm (**l**) and 0.57 μm (**m**).

trimethoxysilane (OTMS) or 1*H*,1*H*,2*H*,2*H*-perfluorooctyltriethoxysilane (PFTS) (Supplementary Methods; spectroscopic evidence of silanization is shown in Supplementary Information section 2 and Supplementary Fig. 9c), they form stable dispersions in heptane and hexane (Supplementary Information section 2 and Supplementary Fig. 6a, b).

Surprisingly, highly corrugated OTMS-modified hydrophobic hedgehog particles (OTMS-HPs) also form dispersions in water (Fig. 1g, h, l), and hydrophilic hedgehog particles disperse in representative hydrophobic solvents such as heptane, hexane and toluene (Fig. 4a–h). This illustrates that surface topography can be used to modulate the interaction between microscale particles and disperse them in phobic solvents.

Immediately on sonicating various hydrophobic OTMS-HP and PFTS-modified hydrophobic hedgehog particle (PFTS-HP) formulations in water (Supplementary Information section 2 and Supplementary Fig. 7), we observed the formation of a precipitate on the bottom of the vial, floating aggregates on top of the liquid, and a stable opalescent dispersion. Dispersions remain stable and free of aggregation for a subset of particles for at least five days, as verified by scanning electron microscopy (SEM) (Fig. 1i) and dynamic light scattering (Supplementary Information section 2 and Supplementary Table 1). The percentage of hydrophobic hedgehog particle aggregates floating on the surface of the dispersions increased with elongation of the nanospikes (Supplementary Information section 2 and Supplementary Fig. 7a–d), but the colloidal stability of the particles dispersed in water was also enhanced (Fig. 2k–m).

To exclude the possibility that the observed behaviour arises because our samples contain a subpopulation of hydrophilic OTMS-HPs or PFTS-HPs or represent a special case of Janus colloids, that is, colloids

consisting of distinct hydrophobic and hydrophilic interfacial sectors^{7–9}, we directly probed the hydrophobic nature of the particles after processing them into dried thin films. The filtrate of suspended OTMS-HPs exhibited high water repellency causing the droplets to roll off (the ‘lotus effect’; Supplementary Information section 2, Supplementary Videos 1–3 and Supplementary Fig. 8). Further evidence is obtained by injecting hydrophobic cadmium selenide (CdSe) nanoparticles into an aqueous dispersion of OTMS-HPs: confocal and transmission electron microscopy (TEM) images show the expected anchoring of hydrophobic nanoparticles on the spikes (Fig. 1j, k, Supplementary Information section 2 and Supplementary Fig. 9d), thus confirming their hydrophobicity and the uniformity of surface derivatization. The stability of the hydrophobic hedgehog particles in aqueous dispersion did not change on CdSe adsorption.

The wetting of corrugated surfaces^{4,10–12} is often attributed to a Cassie–Baxter wetting mode^{13,14} and in our case could include formation of an air shell in the vicinity of the μPS core. Such trapped air bubbles¹⁵ might provide buoyancy to the hedgehog particles, but are known to be thermodynamically unstable^{16,17}. The presence of trapped air is verified by adding ethanol and observing gas evolving from the dispersion (Supplementary Video 5), and by observing, in high-resolution confocal microscopy images of the particles, concentric shells with markedly different refractive indices as would be expected if an air shell is present (Fig. 2a, b, Supplementary Information section 3 and Supplementary Fig. 10). Hydrophilic hedgehog particles in water have no air shells, and they appear under the same conditions as uniformly lit particles (Fig. 2c, d, Supplementary Information section 3 and Supplementary Fig. 11).

When adding fluorescent cadmium telluride (CdTe) nanoparticles, stabilized with hydrophilic thioglycolic acid (TGA), to an aqueous dispersion of the hydrophobic hedgehog particles, a dark zone devoid of emission around the hedgehog particles confirms the presence of a layer of air. The dimensions of the emission exclusion zones closely match the diameter of hedgehog particles (Fig. 2e). The fact that similar images were obtained after five days of storage in the dark without agitation (Fig. 2g) attests to the long-term stability of dispersions of our hydrophobic hedgehog particles in water, consistent with long-term trapping of air at macroscale corrugated surfaces¹⁸ (Supplementary Information section 3 and Supplementary Fig. 13). Hydrophilic hedgehog particles in identical luminescent media appear as bright spots with CdTe nanoparticles localized on and between the nanospikes (Fig. 2f, h).

Strong scattering of photons and electrons by ZnO nanospikes prevents successful optical or cryogenic TEM imaging of the air–water interface within hedgehog particles, but we can locate it by taking advantage of the fact that CdTe nanoparticles can self-assemble into nanowires¹⁹ and nanosheets²⁰ at interfaces²¹: a thin layer of CdTe nanoparticles that assembles more than 200 nm in from the ends of the ZnO nanospikes (Fig. 2i, Supplementary Information section 3 and Supplementary Fig. 14) pinpoints the water meniscus (Fig. 2j). This allows us to calculate an average hedgehog particle density of 0.92 g cm^{-3} (Supplementary Information section 3), which closely matches the density of water and explains the buoyancy of the particles.

We must also explain why two individual hydrophobic hedgehog particles do not coalesce on collision. To do so, we refer to the extended Derjaguin–Landau–Verwey–Overbeek (E_{DLVO}) theory, according to which the sum of potentials associated with van der Waals (V_{vdW}), electrical double layer (V_{DL}) and hydrophobic (V_{HB}) interactions approximate the total interaction potential ($V_{E_{\text{DLVO}}}$) between the hydrophobic hedgehog particles: $V_{E_{\text{DLVO}}} = V_{\text{vdW}} + V_{\text{DL}} + V_{\text{HB}}$. Evaluating interparticle interactions in different configurations (Fig. 3a–c, Supplementary Information section 4 and Supplementary Fig. 19), we find that hedgehog-particle/hedgehog-particle pair potentials display high repulsive energy barriers of at least $14k_{\text{B}}T$ (k_{B} , Boltzmann's constant) for the outer contour of spikes ($x = 0$; Fig. 3d). Penetration of the nanospikes into the interstitial spaces of another hedgehog particle ($x < 0$) is energetically unfavourable (Fig. 3d).

Comparison of the $V_{E_{\text{DLVO}}}$ for hydrophobic hedgehog particles with that for hydrophobic μPS s (Supplementary Information section 4 and Supplementary Fig. 24d) shows that the interfacial corrugations transform the overall attractive potential into a repulsive one. For hedgehog particles with short nanospikes, $V_{E_{\text{DLVO}}}$ reverses such that the interaction becomes attractive (Fig. 3e), matching the experimental results in Fig. 2k. The key reason for the anomalous stability of hedgehog particle dispersions is that V_{vdW} and V_{HB} are greatly decreased for corrugated particles compared with the smooth spheres (Fig. 3f, g), owing to the drastic reduction in the contact area in the former case. The total contour area of tapered spikes represents $<3\%$ of the surface area of the smooth particles (Fig. 1a).

The colloidal stability of hydrophobic hedgehog particles in water is also enhanced by the presence of the double electric layer at the air–water interface, increasing V_{DL} . The zeta-potential (ζ) of air bubbles, which affects their electrostatic repulsion, is known to be between -35 mV (ref. 22) and -65 mV (ref. 23). Such high ζ is attributed to autoionization of water²⁴ that may also occur at the hydrophobic interfaces^{24,25}. However, the fact that the hedgehog particle dispersion remains stable in the presence of $0.01\text{--}1.0 \text{ M NaCl}$, which leads to screening of electrostatic interactions and coagulation of ‘normal’ dispersions (Fig. 1m, n), shows that any increased electrostatic repulsion has a secondary role and that the anomalous colloidal behaviour of hedgehog particles is dominated by the reduction of attractive interactions between the particles. But limitations of Derjaguin–Landau–Verwey–Overbeek theory for high ionic strengths and nanoscale corrugated surfaces²⁶ may need to be considered for a more complete mechanistic explanation.

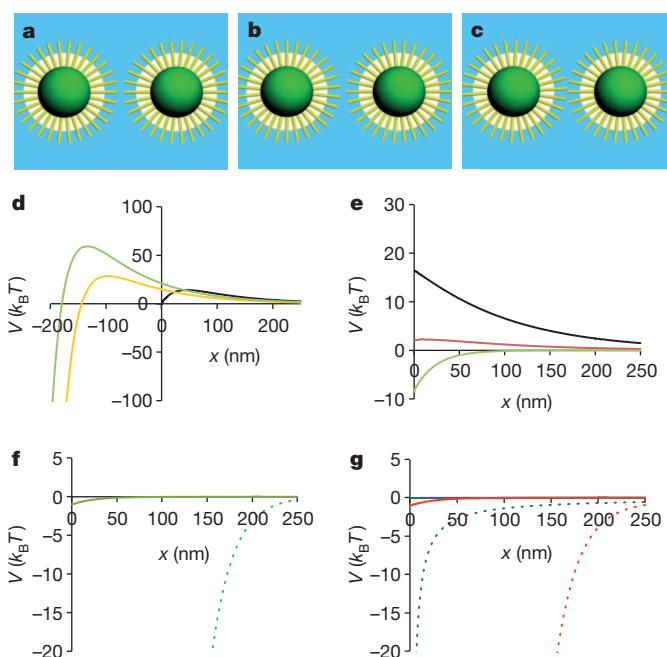


Figure 3 | Interaction potentials of hydrophobic hedgehog particles in aqueous media. **a–c**, Two general configurations, spike-to-spike (S–S; **a**) and spike-to-gap (S–G; **b**), are considered, along with the intermediate case in which the ZnO nanospikes face the side walls of opposing particles (ZS–G; **c**). **d–g**, Interaction potentials between hydrophobic hedgehog particles. **d**, Pair potentials for hydrophobic hedgehog particles in S–S ($V_{E_{\text{DLVO},\text{S-S}}}$, black), S–G ($V_{E_{\text{DLVO},\text{S-G}}}$, orange) and ZS–G ($V_{E_{\text{DLVO},\text{ZS-G}}}$, green) configurations. The negative values of x correspond to the penetration of ZnO nanospikes into the interstitial spaces of another hedgehog particle; $x = 0$ corresponds to the outer contour around the spike tips. **e**, Pair potentials ($V_{E_{\text{DLVO},\text{HP}}}$) of hydrophobic hedgehog particles in an aqueous dispersion calculated according to the E_{DLVO} theory for the zeta-potentials at the air–water interface with $\zeta = -65 \text{ mV}$ (black line) and $\zeta = -35 \text{ mV}$ (red line), and for hydrophobic hedgehog particles with short nanospikes from Fig. 2k (green line). **f**, Hydrophobic interaction potentials of OTMS-HPs ($V_{\text{HB},\text{HP}}$, green) and OTMS- μPS s ($V_{\text{HB},\text{PS}}$, dotted green). **g**, Van der Waals interaction potentials of OTMS-HPs ($V_{\text{vdW},\text{HP}}$, blue) and OTMS- μPS s ($V_{\text{vdW},\text{PS}}$, dotted blue) and total attractive potentials of OTMS-HPs ($V_{\text{vdW}+\text{HB},\text{HP}}$, red) and hydrophobic OTMS- μPS ($V_{\text{vdW}+\text{HB},\text{PS}}$, dotted red) in water.

If the drastic reduction in attractive components of the pair potential is the reason for the unusual stability of hedgehog particle dispersions, the same effect should occur in dispersions of hydrophilic colloids in hydrophobic solvents. Stable dispersions of hydrophilic hedgehog particles were obtained in heptane, hexane and toluene (Fig. 4a, Supplementary Information section 5 and Supplementary Fig. 26). SEM and confocal microscopy images (Fig. 4b–e) demonstrated non-agglomerated particles in the bulk of these dispersions and physical integrity of hedgehog particles (Fig. 4f–h). The μPS core of the hedgehog particles was dissolved in toluene, thus yielding a dispersion of hydrophilic hedgehog particle shells. As expected, ZnO nanoparticles and ZnO nanowires (Supplementary Information section 5 and Supplementary Fig. 27) do not disperse in the same solvents.

Calculations show that V_{vdW} for this type of dispersion is much reduced compared with smooth spheres, and that the overall pair potential of hydrophilic hedgehog particles in heptane is indeed repulsive with $V_{\text{DLVO},\text{HPs}} = 1.4k_{\text{B}}T$ at $x = 0 \text{ nm}$ (Fig. 4i, j, Supplementary Information section 5 and Supplementary Fig. 28). Notably, dispersion in organic solvents lack the air layer between the spikes, and electrostatic interactions in organic solvents are not screened as in aqueous solutions and are therefore longer ranged.

The stability of our surfactant-free hedgehog particles in ‘phobic’ solvents offers a different perspective on scientific and technological problems related to colloidal interactions and might even enable new strategies

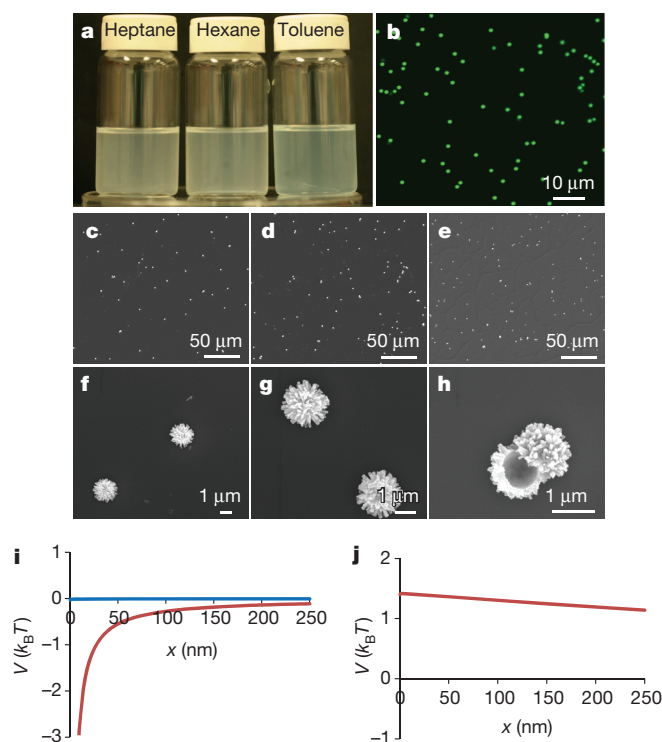


Figure 4 | Dispersion of hydrophilic hedgehog particles in hydrophobic organic solvents. **a**, Dispersions of hydrophilic hedgehog particles in (left to right) heptane, hexane and toluene. As in the case of dispersion of hydrophobic hedgehog particles in water, after sonication of hydrophilic hedgehog particles in organic solvent there was always a small amount of precipitate in the bottom of the vial. **b**, Confocal microscopy image of hydrophilic hedgehog particles in heptane. **c–e**, SEM images of hydrophilic hedgehog particles from dispersions in heptane (**c**), hexane (**d**) and toluene (**e**). **f–h**, SEM images of individual hedgehog particles in heptane (**f**), hexane (**g**) and toluene (**h**). Toluene dissolves the μ PS core in the hedgehog particles, rendering dispersions of hydrophilic spiky shells. **i**, Van der Waals interaction potentials V_{vdW} of hydrophilic hedgehog particles (blue) and μ PS (red) in heptane. **j**, Total pair potential $V_{\text{E_DLVO}}$ of hydrophilic hedgehog particles in heptane.

for processing and dealing with colloids, for developing new drug delivery systems²⁷, and for colloidal self-assembly^{9,28,29}. We also believe that the unusual solvation behaviour of hedgehog particles (contrary to the traditional expectations of particle dispersion stability in hydrophobic/hydrophilic solvents) could be used to develop efficient adsorbers, absorbers, scatterers or catalysts that need to function in both organic and aqueous media.

Received 25 October 2013; accepted 12 November 2014.

1. Israelachvili, J. N. *Intermolecular and Surface Forces* 253–413 (Academic, 2011).
2. Sperling, R. A. & Parak, W. J. Surface modification, functionalization and bioconjugation of colloidal inorganic nanoparticles. *Phil. Trans. R. Soc. A* **368**, 1333–1383 (2010).
3. Feng, L. *et al.* Super-hydrophobic surfaces: from natural to artificial. *Adv. Mater.* **14**, 1857–1860 (2002).
4. Lafuma, A. & Quéré, D. Superhydrophobic states. *Nature Mater.* **2**, 457–460 (2003).
5. Parker, A. R. & Lawrence, C. R. Water capture by a desert beetle. *Nature* **414**, 33–34 (2001).
6. Yang, P. D. *et al.* Controlled growth of ZnO nanowires and their optical properties. *Adv. Funct. Mater.* **12**, 323–331 (2002).
7. Kim, S. H., Abbaspourrad, A. & Weitz, D. A. Amphiphilic crescent-moon-shaped microparticles formed by selective adsorption of colloids. *J. Am. Chem. Soc.* **133**, 5516–5524 (2011).

8. Walther, A. & Muller, A. H. E. Janus particles: synthesis, self-assembly, physical properties, and applications. *Chem. Rev.* **113**, 5194–5261 (2013).
9. Yan, J., Bloom, M., Bae, S. C., Luijten, E. & Granick, S. Linking synchronization to self-assembly using magnetic Janus colloids. *Nature* **491**, 578–581 (2012).
10. Papadopoulos, P., Mammen, L., Deng, X., Vollmer, D. & Butt, H. J. How superhydrophobicity breaks down. *Proc. Natl Acad. Sci. USA* **110**, 3254–3258 (2013).
11. Verho, T. *et al.* Mechanically durable superhydrophobic surfaces. *Adv. Mater.* **23**, 673–678 (2011).
12. Wang, S. & Jiang, L. Definition of superhydrophobic states. *Adv. Mater.* **19**, 3423–3424 (2007).
13. Quéré, D. Wetting and roughness. *Annu. Rev. Mater. Res.* **38**, 71–99 (2008).
14. Whyman, G. & Bormashenko, E. How to make the Cassie wetting state stable? *Langmuir* **27**, 8171–8176 (2011).
15. Tyrrell, J. W. G. & Attard, P. Images of nanobubbles on hydrophobic surfaces and their interactions. *Phys. Rev. Lett.* **87**, 176104 (2001).
16. Mao, M., Zhang, J. H., Yoon, R. H. & Ducker, W. A. Is there a thin film of air at the interface between water and smooth hydrophobic solids? *Langmuir* **20**, 1843–1849 (2004).
17. Martinez, E. *et al.* Air-trapping on biocompatible nanopatterns. *Langmuir* **22**, 11230–11233 (2006).
18. Poetes, R., Holtzmann, K., Franze, K. & Steiner, U. Metastable underwater superhydrophobicity. *Phys. Rev. Lett.* **105**, 166104 (2010).
19. Tang, Z. Y., Kotov, N. A. & Giersig, M. Spontaneous organization of single CdTe nanoparticles into luminescent nanowires. *Science* **297**, 237–240 (2002).
20. Tang, Z. Y., Zhang, Z. L., Wang, Y., Glotzer, S. C. & Kotov, N. A. Self-assembly of CdTe nanocrystals into free-floating sheets. *Science* **314**, 274–278 (2006).
21. Dong, A. G., Chen, J., Vora, P. M., Kikkawa, J. M. & Murray, C. B. Binary nanocrystal superlattice membranes self-assembled at the liquid–air interface. *Nature* **466**, 474–477 (2010).
22. Takahashi, M. Zeta potential of microbubbles in aqueous solutions: electrical properties of the gas–water interface. *J. Phys. Chem. B* **109**, 21858–21864 (2005).
23. Graciaa, A., Morel, G., Saulner, P., Lachaise, J. & Schechter, R. S. The zeta-potential of gas-bubbles. *J. Colloid Interface Sci.* **172**, 131–136 (1995).
24. Beattie, J. K., Djerdjev, A. N. & Warr, G. G. The surface of neat water is basic. *Faraday Discuss.* **141**, 31–39 (2009).
25. Beattie, J. K. & Djerdjev, A. M. The pristine oil/water interface: surfactant-free hydroxide-charged emulsions. *Angew. Chem. Int. Ed.* **43**, 3568–3571 (2004).
26. Huang, R. X., Carney, R. P., Stellacci, F. & Lau, B. L. T. Colloidal stability of self-assembled mono layer-coated gold nanoparticles: the effects of surface compositional and structural heterogeneity. *Langmuir* **29**, 11560–11566 (2013).
27. Saltzman, W. M. & Torchilin, V. *Drug Delivery Systems* (McGraw-Hill, 2008).
28. Zhou, H. *et al.* Self-assembly mechanism of spiky magnetoplasmonic supraparticles. *Adv. Funct. Mater.* **24**, 1439–1448 (2014).
29. Hollingsworth, A., Leunissen, M., Irvine, W., Chaikin, P. & van Blaaderen, A. Charged colloids in low polar solvents. *Bull. Am. Phys. Soc.* **53**, abstr. L9.00012 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements This material is based on work partially supported by the Center for Solar and Thermal Energy Conversion, an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under award number DE-SC0000957. We acknowledge support from the US NSF under grant ECS-0601345, CBET 0933384, CBET 0932823 and CBET 1036672. The work is also partly supported by the US Department of Defense under grant awards nos W911NF-10-1-0518, MURI W911NF-12-1-0407 and MURI W911NF-12-1-0407. We want express our appreciation to N. Walter for his leadership in organizing the SMART Center at the University of Michigan; J. Young Kim, Q. Che and X. Lu for help with the synthesis of CdTe nanoparticles; D. Bukharina for help with synthesis of hedgehog particles and SEM imaging of samples for dispersions in hydrophobic solvents; and D. Sorenson for assistance with TEM imaging of nanoparticle assemblies on OTMS-HPs. We also want to express our appreciation of H. Yu and G. Khan, who calculated van der Waals interactions of ZnO nanorods using the coupled dipole method. The guidance of C. Silvera Batista in this process is also much appreciated.

Author Contributions N.K. had the idea for the project. N.K. and J.H.B. designed the experiments. J.H.B. designed and made the hedgehog particles, did the experiments and did the calculations of the pairwise interaction potential between the hedgehog particles. J.H.B. and B.Y. did the initial design of the hedgehog particles. Y.W. did biological experiments and helped to make the hedgehog particles. S.O.T. made the Fourier transform infrared spectroscopy measurements and helped to make the hedgehog particles. J.D.H. did confocal microscopy using an Olympus IX81 inverted with an ISS Alba 5 to detect concentric mesoscale shells on hedgehog particles indicating differences in refractive indices.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to N.K. (kotov@umich.edu).

Metal-catalysed azidation of tertiary C–H bonds suitable for late-stage functionalization

Ankit Sharma¹ & John F. Hartwig¹

Many enzymes oxidize unactivated aliphatic C–H bonds selectively to form alcohols; however, biological systems do not possess enzymes that catalyse the analogous aminations of C–H bonds^{1,2}. The absence of such enzymes limits the discovery of potential medicinal candidates because nitrogen-containing groups are crucial to the biological activity of therapeutic agents and clinically useful natural products. In one prominent example illustrating the importance of incorporating nitrogen-based functionality, the conversion of the ketone of erythromycin to the –N(Me)CH₂– group in azithromycin leads to a compound that can be dosed once daily with a shorter treatment time^{3,4}. For such reasons, synthetic chemists have sought catalysts that directly convert C–H bonds to C–N bonds. Most currently used catalysts for C–H bond amination are ill suited to the intermolecular functionalization of complex molecules because they require excess substrate or directing groups, harsh reaction conditions, weak or acidic C–H bonds, or reagents containing specialized groups on the nitrogen atom^{5–14}. Among C–H bond amination reactions, those forming a C–N bond at a tertiary alkyl group would be particularly valuable, because this linkage is difficult to form from ketones or alcohols that might be created in a biosynthetic pathway by oxidation¹⁵. Here we report a mild, selective, iron-catalysed azidation of tertiary C–H bonds that occurs without excess of the valuable substrate. The reaction tolerates aqueous environments and is suitable for the functionalization of complex structures in the late stages of a multistep synthesis. Moreover, this azidation makes it possible to install a range of nitrogen-based functional groups, including those from Huisgen ‘click’ cycloadditions and the Staudinger ligation^{16–19}. We anticipate that these reactions will create opportunities to modify natural

products, their precursors and their derivatives to produce analogues that contain different polarity and charge as a result of nitrogen-containing groups. It could also be used to help identify targets of biologically active molecules by creating a point of attachment—for example, to fluorescent tags or ‘handles’ for affinity chromatography—directly on complex molecular structures.

To develop a mild method for the conversion of an alkyl C–H bond to an alkyl C–N bond, we focused on reactions of the hypervalent iodine reagent **1** containing an azide unit (Fig. 1). Such a reagent is related to hypervalent reagents commonly used for oxidation²⁰ and is thermally stable (up to 130°C)²¹. It has sufficient thermodynamic potential to convert alkyl C–H bonds to alkyl azides, but the published reactions have been limited to simple hydrocarbons, typically used in excess amounts, or activated C–H bonds at high temperatures in the presence of radical initiators²¹. Thus, the current azidations of C–H bonds by this reagent²¹ are not suitable for late-stage functionalization of complex molecules. If an appropriate transition-metal catalyst for C–H bond functionalization with this hypervalent iodine reagent could be identified, then C–H bond amination reactions that incorporate an azide into complex molecules with site selectivity could be devised. Previously, iron- and manganese-porphyrin complexes were reported to catalyse the formation of alkyl azides using sodium azide and iodosobenzene or *t*BuOOH as oxidants, but the reactions were limited to hydrocarbons, an excess of the alkane (10 equiv.) was required, and a major by-product was the corresponding alcohol^{22–24}.

To identify a metal complex that would catalyse the azidation of C–H bonds with **1** under mild conditions, we investigated the reaction of *cis*-decalin. Various metal complexes, including metal porphyrins, were

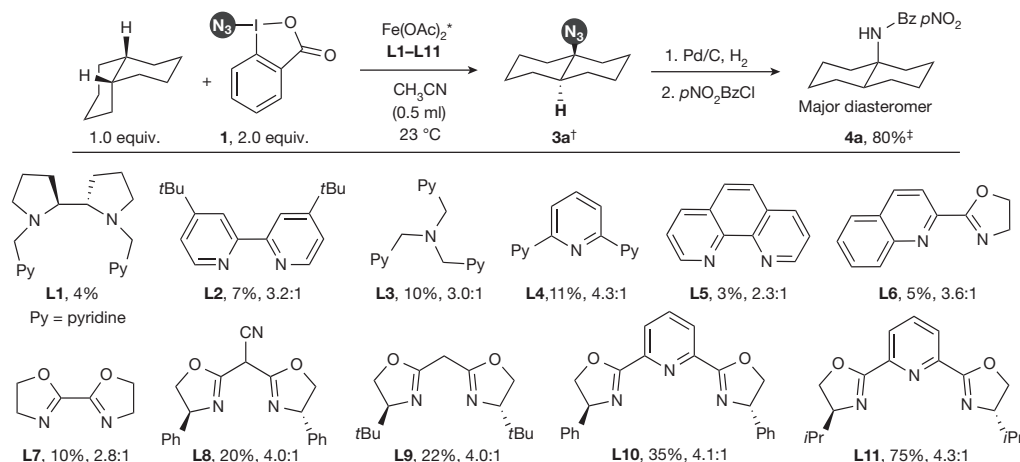


Figure 1 | Development of a catalyst for the azidation of aliphatic C–H bonds. Top row, reactions studied. *Conditions: 10.0 mol% Fe(OAc)₂, 11.0 mol% ligand (L1–L11), *cis*-decalin (0.2 mmol, 1.0 equiv.) and **1** (0.4 mmol, 2.0 equiv.), 23°C, 1 ml of CH₃CN. †The relative configuration of the major diastereomer of **3a** was confirmed by X-ray crystallographic analysis of **4a**.

[‡]Combined yield of isomers. Middle and bottom rows, structural formulae of ligands L1–L11. Also shown are yields and the ratios of isomers determined by gas chromatography analysis with dodecane as internal standard; ratios are not corrected for response factors of minor isomers.

¹Department of Chemistry, University of California, Berkeley, California 94720, USA.

tested as catalysts with the hydrocarbon as limiting reagent; only iron complexes provided measurable yields of azide product at 23 °C (Supplementary Table 1). Combinations of $\text{Fe}(\text{OAc})_2$ and various bi-, tri- and tetra-dentate nitrogen ligands (**L1**–**L5** in Fig. 1), including those that catalyse the selective hydroxylation of aliphatic C–H bonds (**L1** and **L3**)^{25,26}, provided low yields of the product **3a** (Fig. 1; 3–11%). However, substantial yields of **3a** were observed with iron complexes of oxazoline-derived ligands, particularly with those possessing larger N–Fe–N ‘bite’ angles (**L6**–**L9** in Fig. 1). Finally, we found that reactions of iron complexes containing tridentate nitrogen ligands of the pybox family (**L10** and **L11** in Fig. 1) provided product **3a** in good yield with high selectivity for reaction at a tertiary C–H bond (75%, 4.3:1 ratio of diastereomers). Reactions conducted in ethyl acetate (EtOAc)

with the catalyst containing ligand **L11** proceeded in good yield and higher *trans:cis* selectivity (65%, 6.3:1; Supplementary Table 2). Reactions in acetonitrile were faster, but occurred with lower selectivity. Reactions in a mixture of EtOAc and water (5:1) occurred similarly to those in pure EtOAc (63%, 6:1).

The azidation of the C–H bond of a series of hydrocarbons occurred in excellent yields with high selectivity for a tertiary C–H bond over the secondary and primary C–H bonds (Supplementary Table 3), setting the stage for azidation of the tertiary C–H bonds in molecules containing a series of functional groups. The azidation reaction with derivatives of dihydrocitronellol containing two electronically distinct tertiary C–H bonds and many secondary C–H bonds is shown in Fig. 2a. These reactions revealed the inherent electronic selectivity of the azidation reaction

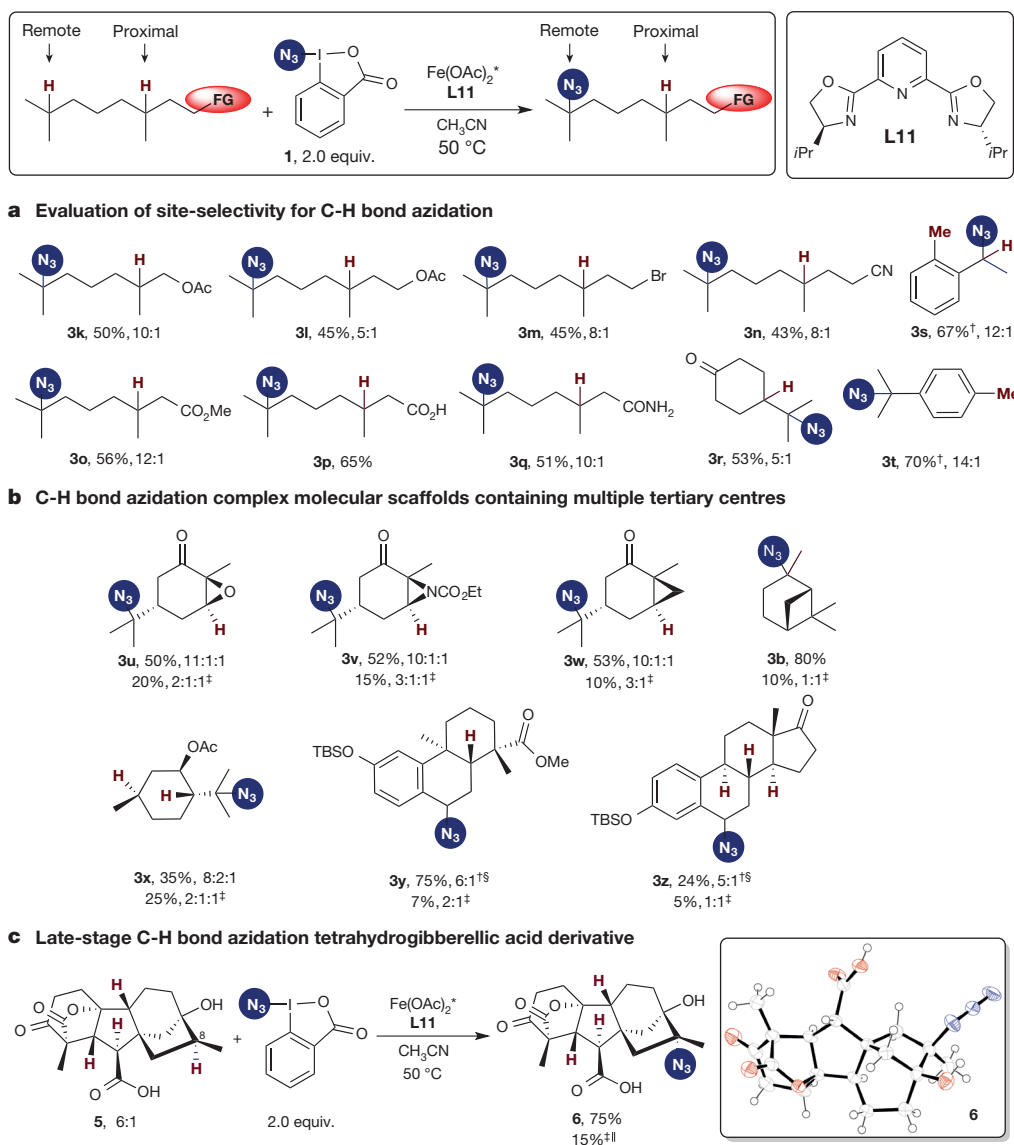


Figure 2 | Evaluation of the effect of the steric and electronic environment on site selectivity for the azidation of aliphatic tertiary C–H bonds. Top row, reaction studied. *Conditions: 10.0 mol% $\text{Fe}(\text{OAc})_2$, 11.0 mol% ligand **L11**, substrate (on left: 0.2 mmol, 1.0 equiv.) and **1** (0.6 mmol, 3.0 equiv.), 23 °C, 1 ml of CH_3CN . Isolated yields of major azide products are reported unless mentioned otherwise. The ratios of isomers were determined by gas chromatography analysis with dodecane as internal standard, and are not corrected for response factors of minor isomers. **a**, Evaluation of site selectivity for C–H bond azidation. The ratios reported reflect the site selectivity for reaction at the two tertiary C–H bonds (**3k** to **3r**) or the two benzylic C–H bonds (**3s**, **3t**). **3k**–**3t** are the products formed from the azidation. [†]EtOAc was

used as solvent. **b**, C–H bond azidation of complex molecular scaffolds containing multiple tertiary centres. Yields of products **3u**–**3z** were also compared to those of the benzoyl-peroxide-initiated reaction, as follows.

[‡]Conditions: 10.0 mol% BzOObz , substrate (0.2 mmol, 1.0 equiv.) and **1** (0.6 mmol, 3.0 equiv.), 84 °C and 1,2-dichloroethane (1.0 ml) as solvent. The yield and ratios of isomers were determined by gas chromatography analysis with dodecane as internal standard. The ratios refer to the major product versus minor product determined to be isomers by mass spectrometry. **c**, Late-stage C–H bond azidation of a tetrahydrogibberellic acid derivative. As reaction at top, except substrate is **5**, and **6** is product. [§]Diastereoselectivity was measured by ¹H NMR of crude reaction mixture. ^{||}Unidentifiable mixture of products.

and its functional-group compatibility. The C–H bond azidation was selective for reaction at the more electron-rich, remote, tertiary C–H bond, resulting in good isolated yields of the pure major isomers formed by the reaction (**3k–3q** in Fig. 2a). The regioselectivity of azidation at the two electronically distinct tertiary C–H bonds was influenced by the distance of the electron-withdrawing group from the proximal tertiary C–H bond (**3k** and **3l**). In these cases, the regioselectivity of the C–H bond azidation reaction mirrors the regioselectivity of a wide range of oxidation reactions²⁷. Functional groups—such as an alcohol protected as an acetoxy group (**3k** and **3l**), a bromide (**3m**), a nitrile (**3m**), an ester (**3o**), a carboxylic acid (**3p**) and an amide (**3q**)—were tolerated. Functional groups like a carboxylic acid (**3p**) and an amide (**3q**) that could act as directing groups influenced the selectivity by their electronic properties, rather than by coordination to the catalyst. This higher reactivity of more electron-rich C–H bonds was also observed for cyclic structures. 4-*iso*-Propylcyclohexanone underwent azidation with high regioselectivity for the more electron-rich of the two tertiary C–H bonds (**3r** in Fig. 2a). Investigation of the reactions of substituted arenes showed that tertiary and secondary benzylic C–H bonds were functionalized selectively in the presence of primary benzylic C–H bonds (**3s** and **3t** in Fig. 2a).

Having revealed high regioselectivity for C–H bond azidation, we assessed the potential of this reaction for azidation of the C–H bonds in more complex scaffolds containing several functional groups and strained rings that could react instead of a C–H bond or influence the identity of the C–H bond that undergoes azidation (Fig. 2b). Cyclic ketones prepared from (–)-carvone underwent azidation at the tertiary C–H bond remote from the ketone with high regioselectivity. These reactions occurred in the presence of epoxides, aziridines and cyclopropanes in good isolated yields (**3u–3w** in Fig. 2b). Minor azidation products were also observed by gas chromatography–mass spectrometry. These products were formed in amounts too small for isolation and were not characterized. A mixture of diastereomers of α -dihydropinene (**2b**, 5:1) containing three electronically similar, but sterically distinct, tertiary C–H bonds reacted to give 80% isolated yield of a single isomer of azide **3b** at room temperature. The strained four-membered ring was tolerated, suggesting a fast recombination of the likely radical intermediates. Acetoxymethylol containing two electronically similar tertiary C–H bonds reacted preferentially at the *iso*-propyl side chain to provide one major constitutional isomer in moderate isolated yield (**3x** in Fig. 2b). This selectivity, presumably, results from the greater conformational flexibility of the *iso*-propyl side chain. Isomeric products from azidation of a different C–H bond were observed as minor products; again, these products were formed in quantities too low for full characterization, but were shown to be isomers by mass spectrometry.

Unlike the stereoretentive property of the metal-catalysed insertions of nitrenes or carbenes into C–H bonds^{8,9,28}, the configuration of the

carbon bound to the azide is independent of the configuration in the reactant. However, this stereochemical outcome of our reactions allows one to use mixtures of diastereomeric reactants (see Supplementary Table 3 and Fig. 2; explicitly, **2b** to form **3b** and **5** to form **6**) to provide one major diastereomer of the azide product (see below).

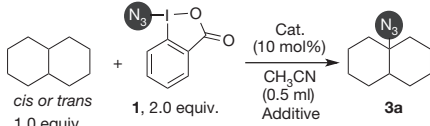
Biologically active molecules containing multiple benzylic and tertiary C–H bonds also reacted selectively. Podocarpic acid and its derivatives have been reported to exhibit a wide variety of biological activities, including antileukaemic activity, inhibition of plant cell growth, and anti-inflammatory properties. A podocarpic acid derivative underwent selective azidation at the benzylic C–H bond in high yields and good diastereoselectivity (**3y** in Fig. 2b). Similar selectivity was also observed for the azidation of an oestrone (**3z** in Fig. 2b).

The reaction of a gibberellic acid derivative illustrates the ability to conduct the azidation of complex structures (Fig. 2c). Gibberellic acid is a plant hormone that regulates growth and influences developmental processes, including cell elongation and germination. The gibberellic acid derivative **5** is a pentacyclic diterpene containing four tertiary C–H bonds. Based on the data just presented, the most electron-rich and sterically least hindered tertiary C–H bond, the one at carbon 8, should react selectively. In addition, the stereochemical outcome of the azidation of *cis*- and *trans*-decalin and α -dihydropinene (Supplementary Table 3 and Fig. 2b) suggested that the configuration of the reactive centre in substrate **5** would have a negligible influence on the diastereomeric ratio of product **6** (Fig. 2c). Indeed, the azidation of a mixture of diastereomers of **5** provided the corresponding azide **6** as a single isolated diastereoisomer in 75% yield (Fig. 2c) from *exo*-attack of the azide unit at C(8).

Finally, we also tested functionalizations of the complex scaffolds shown in Fig. 2b and c by reactions initiated with benzoyl peroxide. In all cases, poor yields and selectivities were observed from the reactions initiated by the peroxide. The yields of the products from reaction of the substrates in Fig. 2b were low in all cases and formed mixtures of isomeric products with poor selectivity. In addition, gibberellic acid derivative **5** decomposed to form a complex mixture of products in the presence of azide **1** and the peroxide. This distinct reaction course in the presence and absence of the iron catalyst suggests that the C–N bond is formed by two different processes in the two systems, and underscores the importance of the iron catalyst to create a reaction that is suitable for late-stage functionalization of complex molecules.

Although detailed mechanistic studies have not yet been conducted, several observations reveal the general features of the mechanism. The site selectivities and stereochemical outcome of the azidation of *cis*- and *trans*-decalin and α -dihydropinene strongly suggest that a tertiary alkyl radical is generated (Supplementary Table 3 and Fig. 2b). Attempts to use radical clocks to assess more directly a potential alkyl radical were

Table 1 | Experiments to evaluate the involvement of radical intermediates and the role of the iron catalyst

Entry	Substrate	Catalyst	Temperature (°C)	Additive	Yield (%)	Selectivity
1	<i>cis</i> or <i>trans</i> 1.0 equiv.		23	TEMPO*	3	NA
2	<i>cis</i>	Fe(OAc) ₂ /L11	23	BHT*	3	NA
3†	<i>cis</i>	Fe(OAc) ₂ /L11	80	NA	55	3.2
4†	<i>trans</i>	Fe(OAc) ₂ /L11	80	NA	43	3.2
5†	<i>cis</i>	BzOOBz	80	ABCN‡	40	1.7
6†	<i>trans</i>	BzOOBz	80	ABCN‡	33	1.7

Conditions: 10.0 mol% catalyst, *cis*- or *trans*-decalin (0.2 mmol, 1.0 equiv.) and **1** (0.4 mmol, 2.0 equiv.), 2 h. The yield and ratios of isomers were determined by gas chromatography analysis with dodecane as internal standard and not corrected for response factors of minor isomers. NA, not applicable.

* 1.0 equiv. was added.

† EtOAc was used as solvent.

‡ 1.0 mol% was added.

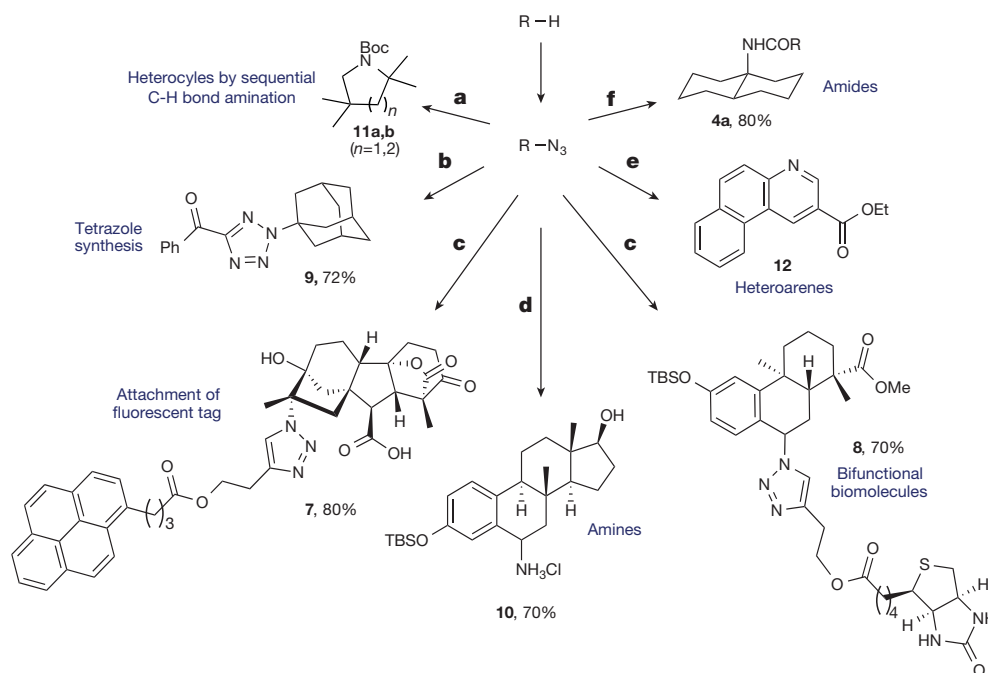


Figure 3 | Introduction of a series of nitrogen-containing functionalities via C–H bond azidation. **a**, Azide (1.0 equiv.), Fe cat., Fmoc-OSuc (1.5 equiv.), 65°C, benzene, 24 h (see ref. 19 for details). **b**, BzCN (2 equiv.), 130°C, 48 h; **c**, CuSO₄ (10 mol%), alkyne (2 equiv.), DMF, 48 h; **d**, CuSO₄ (10 mol%),

NaBH₄ (3 equiv.), MeOH; **e**, (1) TfOH (1.0 equiv.), toluene (2.0 equiv.), ethyl 3-ethoxyacrylate; (2) 2,3-Dichloro-5,6-dicyano-1,4-benzoquinone (DDQ), EtOAc, 5 min (ref. 18); **f**, (1) Pd/C, H₂, MeOH; (2) Ac₂O, dichloromethane, 12 h.

hampered by the poor reactivity of the appropriate substrates (see Supplementary Information for more details), but the proposed radical intermediate is consistent with the selectivity for azidation of the more electron-rich, less polarized, and thus weaker, tertiary C–H bonds (Fig. 2)²⁷. Furthermore, addition of 1 equiv. of BHT and TEMPO (structures shown in Table 1), which are known to quench radicals, resulted in complete inhibition of the azidation reaction (Table 1, entries 1 and 2). Finally, the kinetic isotope effect (KIE) for azidation of ethylbenzene and ethylbenzene-*d*₁₀ in separate vessels from initial reaction rates was observed to be 5.0 ± 0.3 , implying that the cleavage of the C–H bond is the overall turnover-limiting step.

To assess the role of the iron catalyst in this transformation, we compared the iron-catalysed azidations of the complex scaffolds in Fig. 2b and c with the reactions initiated by benzoyl peroxide. As noted above, poor yields were observed in all cases from the reactions initiated by the peroxide. The selectivities for formation of the azidation product from these reactions were lower than those of the iron-catalysed reactions. In addition, the diastereomeric ratio of product **3a** formed from decalin and azide **1** in the presence of an organic radical initiator was different from that formed from the iron-catalysed reaction conducted at the 80°C required for the peroxide-initiated process (Table 1, entries 3–6). These differences in selectivities are all consistent with a different species forming the C–N bond during the iron-catalysed reaction and during the radical-initiated process. One possible origin of this difference is formation of the C–N bond in the iron-catalysed process by reaction of an alkyl radical with an iron azide intermediate.

This C–H bond azidation creates access to a range of synthetically useful functionalities attached to the original substrate by a C–N bond (Fig. 3)^{16–19}. The primary amine formed from azides **3a** and **3z** containing a fully or partially substituted carbon atom (**4a** and **10**), and heterocycles such as tetrazole **9** from azide **3f**, form in good yields (Supplementary Table 3)²⁹. The azides (for example, **3d** and **3e**, Supplementary Table 3) also undergo intramolecular cyclization under conditions reported recently¹⁹ to form various heterocycles, creating a route to nitrogen heterocycles, such as **11**, from alkanes by two C–H bond amination reactions. Finally, the azide functionality undergoes Huisgen cycloaddition reactions. For example, an alkyne tethered to a fluorescent tag coupled

with azido gibberellic acid derivative **6**, and an alkyne attached to biotin coupled with azido podocarpic acid derivative **3y**. These reactions illustrate how azidation and cycloaddition can create bioconjugation methods for visualization and identification of cellular targets of biologically active natural products³⁰.

Much development of this C–H bond functionalization method remains to be accomplished, but a wide range of applications and extensions of the azidation reaction can be envisioned. The modularity of the catalyst creates further opportunities for site selectivity, and the stereochemical content of the ligand creates the potential for enantioselective azidation. The cycloadditions of azides could make possible conjugation to antibodies, and the simple reduction of the azide and the tolerance of the reaction to water creates the potential to intercept biosynthetic sequences and install an amino group in place of a hydroxyl group in the final stages. Finally, we anticipate that this process will spur development of new classes of catalysts for the azidation of C–H bonds that could proceed by distinct mechanisms with distinct selectivities for primary, secondary and tertiary C–H bonds. As rhodium-catalysed amination reactions develop further, the two classes of systems for C–H bond amination should begin to provide a set of tools for incorporation of nitrogen atoms that parallels the existing set of tools for the chemical and enzymatic oxidation of C–H bonds.

Received 8 October; accepted 28 November 2014.

- Bollinger, J. M. & Broderick, J. B. Frontiers in enzymatic C–H-bond activation. *Curr. Opin. Chem. Biol.* **13**, 51–57 (2009).
- Lewis, J. C., Coelho, P. S. & Arnold, F. H. Enzymatic functionalization of carbon-hydrogen bonds. *Chem. Soc. Rev.* **40**, 2003–2021 (2011).
- Plouffe, J. et al. Clinical efficacy of intravenous followed by oral azithromycin monotherapy in hospitalized patients with community-acquired pneumonia. *Antimicrob. Agents Chemother.* **44**, 1796–1802 (2000).
- O'Doherty, B., Muller, O. & Grp, A. S. Randomized, multicentre study of the efficacy and tolerance of azithromycin versus clarithromycin in the treatment of adults with mild to moderate community-acquired pneumonia. *Eur. J. Clin. Microbiol. Infect. Dis.* **17**, 828–833 (1998).
- Jeffrey, J. L. & Sarpong, R. Intramolecular C(sp³)–H amination. *Chem. Sci.* **4**, 4092–4106 (2013).
- Ochiai, M., Miyamoto, K., Kaneaki, T., Hayashi, S. & Nakanishi, W. Highly regioselective amination of unactivated alkanes by hypervalent sulfonylimino- Δ^3 -bromane. *Science* **332**, 448–451 (2011).

7. Louillat, M. L. & Patureau, F. W. Oxidative C-H amination reactions. *Chem. Soc. Rev.* **43**, 901–910 (2014).
8. Roizen, J. L., Harvey, M. E. & Du Bois, J. Metal-catalyzed nitrogen-atom transfer methods for the oxidation of aliphatic C-H bonds. *Acc. Chem. Res.* **45**, 911–922 (2012).
9. Dequierez, G., Pons, V. & Dauban, P. Nitrene chemistry in organic synthesis: still in its infancy? *Angew. Chem. Int. Edn Engl.* **51**, 7384–7395 (2012).
10. Lescot, C., Darses, B., Collet, F., Retaillieu, P. & Dauban, P. Intermolecular C-H amination of complex molecules: insights into the factors governing the selectivity. *J. Org. Chem.* **77**, 7232–7240 (2012).
11. Roizen, J. L., Zalatan, D. N. & Du Bois, J. Selective intermolecular amination of C-H bonds at tertiary carbon centers. *Angew. Chem. Int. Edn Engl.* **52**, 11343–11346 (2013).
12. Michaudel, Q., Thevenet, D. & Baran, P. S. Intermolecular Ritter-type C-H amination of unactivated sp^3 carbons. *J. Am. Chem. Soc.* **134**, 2547–2550 (2012).
13. Li, J. *et al.* Simultaneous structure-activity studies and arming of natural products by C-H amination reveal cellular targets of eupalmerin acetate. *Nature Chem.* **5**, 510–517 (2013).
14. McNally, A., Haffemayer, B., Collins, B. S. L. & Gaunt, M. J. Palladium-catalysed C-H activation of aliphatic amines to give strained nitrogen heterocycles. *Nature* **510**, 129–133 (2014).
15. Pronin, S. V., Reiher, C. A. & Shenvi, R. A. Stereoinversion of tertiary alcohols to tertiary-alkyl isonitriles and amines. *Nature* **501**, 195–199 (2013).
16. Bräse, S., Gil, C., Knepper, K. & Zimmermann, V. Organic azides: an exploding diversity of a unique class of compounds. *Angew. Chem. Int. Edn Engl.* **44**, 5188–5240 (2005).
17. Schilling, C. I., Jung, N., Biskup, M., Schepers, U. & Bräse, S. Bioconjugation via azide-Staudinger ligation: an overview. *Chem. Soc. Rev.* **40**, 4840–4871 (2011).
18. Lallana, E., Riguera, R. & Fernandez-Megia, E. Reliable and efficient procedures for the conjugation of biomolecules through Huisgen azide-alkyne cycloadditions. *Angew. Chem. Int. Edn Engl.* **50**, 8794–8804 (2011).
19. Hennessy, E. T. & Betley, T. A. Complex N-heterocycle synthesis via iron-catalyzed, direct C-H bond amination. *Science* **340**, 591–595 (2013).
20. Zhdankin, V. V. *Hypervalent Iodine Chemistry: Preparation, Structure and Synthetic Applications of Polyvalent Iodine Compounds* (Wiley & Sons, 2013).
21. Zhdankin, V. V. *et al.* Preparation, X-ray crystal structure, and chemistry of stable azidoiodinanes — Derivatives of benziodoxole. *J. Am. Chem. Soc.* **118**, 5192–5197 (1996).
22. Hill, C. L., Smegal, J. A. & Henly, T. J. Catalytic replacement of unactivated alkane carbon-hydrogen bonds with carbon-X bonds (X = nitrogen, oxygen, chlorine, bromine, or iodine) — Coupling of intermolecular hydrocarbon activation by Mn^{III} TPPX complexes with phase-transfer catalysis. *J. Org. Chem.* **48**, 3277–3281 (1983).
23. Kojima, T., Leising, R. A., Yan, S. P. & Que, L. Alkane functionalization at nonheme iron centers. Stoichiometric transfer of metal-bound ligands to alkane. *J. Am. Chem. Soc.* **115**, 11328–11335 (1993).
24. Costas, M. Selective C-H oxidation catalyzed by metalloporphyrins. *Coord. Chem. Rev.* **255**, 2912–2932 (2011).
25. Kim, C., Chen, K., Kim, J. H. & Que, L. Stereospecific alkane hydroxylation with H_2O_2 catalyzed by an iron(II)-tris(2-pyridylmethyl)amine complex. *J. Am. Chem. Soc.* **119**, 5964–5965 (1997).
26. Chen, M. S. & White, M. C. A predictably selective aliphatic C-H oxidation reaction for complex molecule synthesis. *Science* **318**, 783–787 (2007).
27. Newhouse, T. & Baran, P. S. If C-H bonds could talk: selective C-H bond oxidation. *Angew. Chem. Int. Edn Engl.* **50**, 3362–3374 (2011).
28. Davies, H. M. L. & Morton, D. Guiding principles for site selective and stereoselective intermolecular C-H functionalization by donor/acceptor rhodium carbenes. *Chem. Soc. Rev.* **40**, 1857–1869 (2011).
29. Demko, Z. P. & Sharpless, K. B. A click chemistry approach to tetrazoles by Huisgen 1,3-dipolar cycloaddition: synthesis of 5-acyltetrazoles from azides and acyl cyanides. *Angew. Chem. Int. Edn Engl.* **41**, 2113–2116 (2002).
30. Trippier, P. C. Synthetic strategies for the biotinylation of bioactive small molecules. *ChemMedChem* **8**, 190–203 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the US NIH (4R37GM055382 to J.F.H.) and the Swiss National Science Foundation (SNSF; PBGE2_145544 to AS) for financial support. We thank A. DiPasquale for assistance with crystallographic data and acknowledge US NIH shared instrumentation grant S10-RR027172.

Author Contributions A.S. conducted the experiments. A.S. and J.F.H. conceived and designed the project, analysed the data and prepared this manuscript. X-ray crystal structures are deposited in the Cambridge Crystallographic Data Centre (CCDC 1027821–1027822).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.F.H. (jhartwig@berkeley.edu).

Effects of electron correlations on transport properties of iron at Earth's core conditions

Peng Zhang¹, R. E. Cohen^{1,2} & K. Haule³

Earth's magnetic field has been thought to arise from thermal convection of molten iron alloy in the outer core, but recent density functional theory calculations have suggested that the conductivity of iron is too high to support thermal convection^{1–4}, resulting in the investigation of chemically driven convection^{5,6}. These calculations for resistivity were based on electron–phonon scattering. Here we apply self-consistent density functional theory plus dynamical mean-field theory (DFT + DMFT)⁷ to iron and find that at high temperatures electron–electron scattering is comparable to the electron–phonon scattering, bringing theory into agreement with experiments and solving the transport problem in Earth's core. The conventional thermal dynamo picture is safe. We find that electron–electron scattering of *d* electrons is important at high temperatures in transition metals, in contrast to textbook analyses since Mott^{8,9}, and that 4*s* electron contributions to transport are negligible, in contrast to numerous models used for over fifty years. The DFT + DMFT method should be applicable to other high-temperature systems where electron correlations are important.

Recent DFT calculations by Pozzo *et al.*³ predict the electrical resistivity of iron to be $(6.3\text{--}7.5) \times 10^{-5} \Omega \text{ cm}$ at temperatures from 4,580 K to 6,400 K and pressures from 120 GPa to 340 GPa. The thermal conductivities they predicted are approximately three times the currently used values of $46\text{--}63 \text{ W m}^{-1} \text{ K}^{-1}$ in geophysics¹⁰. The results of Pozzo *et al.*³ are consistent with previous DFT studies^{1,2,4,11}. The large electrical and thermal conductivities, however, challenge current Earth models.

Efforts to constrain the transport properties of iron at core conditions have a long history. Elsasser estimated the resistivity of iron to be $\rho \approx 10.0 \times 10^{-5} \Omega \text{ cm}$ at core conditions on the basis of geophysical arguments¹². By assuming the resistivity of iron to be constant along the melting line, Stacey and Anderson obtained $\rho = 11.2 \times 10^{-5} \Omega \text{ cm}$ at 4,971 K and 330 GPa (ref. 10).

All previous calculations neglect electron–electron scattering. It has long been believed that resistivity in ordinary metals arises primarily from electron–phonon scattering, except at cryogenic conditions⁹. Calculations of resistivity from electron–electron scattering only now have become possible owing to developments in computational theory and technology and access to large-scale computational resources. The DFT + DMFT approach has proved successful in providing results that are in good agreement with experiments for iron-bearing compounds^{13,14} and other strongly correlated materials. It quantitatively predicts properties such as magnetic moments and the effective mass of a series of compounds in iron pnictides and iron chalcogenides. It also explains why superconducting gaps in these compounds are strongly Fermi-surface dependent.

Our primary interest is in the properties of Earth's core, so we first present resistivities at the core density of iron (throughout we refer to Earth's core density from seismology of 13.04 g cm^{-3} , or an atomic volume of 47.8 atomic units = 7.083 \AA^3) (Fig. 1). The resistivities calculated by Sha and Cohen¹, de Koker *et al.*² and Pozzo *et al.*³ at the core conditions are approximately half the value obtained by extrapolating from the systematics of Stacey and Anderson¹⁰ and half the value obtained by

extrapolating from previous shock compression experimental results^{15–17}. The thermal conductivity *k* of pure iron at core conditions obtained from their³ calculations ranges from $150 \text{ W m}^{-1} \text{ K}^{-1}$ to about $250 \text{ W m}^{-1} \text{ K}^{-1}$. Assuming a large thermal conductivity, the calculated heat conduction down the core adiabat is about 15 terawatts (TW)³, which overlaps the estimates^{18,19} of total heat loss from the core of 8–16 TW. No energy is left to drive the thermal convection in the geodynamo. To sustain the geodynamo, compositional convection is therefore required^{3,5}. However, this mechanism leads to a new paradox:⁶ Earth's inner core solidification is believed to have started about one billion years ago⁴, so before that there would be no compositional convection to drive the dynamo, yet we know that Earth's geodynamo has existed for more than 3.4 billion years²⁰.

We find that at high temperatures the resistivity from electron–electron scattering, ρ_{ee} , computed with DFT + DMFT is of the same order as the electron–phonon scattering, ρ_{ep} , computed with DFT (Fig. 1). The sum of the two parts of the resistivity, from the electron–electron and the electron–phonon scattering, is in agreement with earlier geophysical estimates. After including both the electron–electron scattering and the electron–phonon scattering, traditional resistivity values are recovered. We checked the systematics of Stacey and Anderson using resistivity results at other density as well. Considering the uncertainty of iron's melting temperature ($\sim \pm 500 \text{ K}$) the resistivity of iron is around $13.5 \times 10^{-5} \Omega \text{ cm}$ along its melting line. Our results support Stacey and Anderson's systematics.

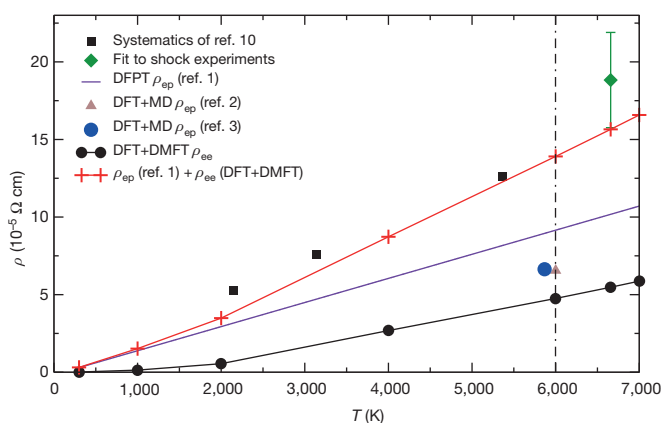


Figure 1 | Resistivity versus temperature of hcp iron at Earth's core density. The black vertical line indicates Earth's core temperature^{30,31}. The black squares are the extrapolations to this density using the systematics of Stacey and Anderson¹⁰. The green diamond is an interpolation to this density of previous shock compression results^{15–17}. The DFPT resistivity line is from the linear extrapolation of low-temperature results^{1,32}. The DFT + molecular dynamics (MD) resistivities are extracted from refs 2,3. The statistical error bars of DFT + DMFT and the total resistivities are smaller than their symbols. Values are given in the Extended Data Tables 1 and 2. All error bars are 1σ .

¹Geophysical Laboratory, Carnegie Institution of Washington, 5251 Broad Branch Road NW, Washington DC 20015, USA. ²Department of Earth Sciences, University College London, 222 Pearson Building, Gower Street, London WC1E 6BT, UK. ³Department of Physics, Rutgers University, Piscataway, New Jersey 08854, USA.

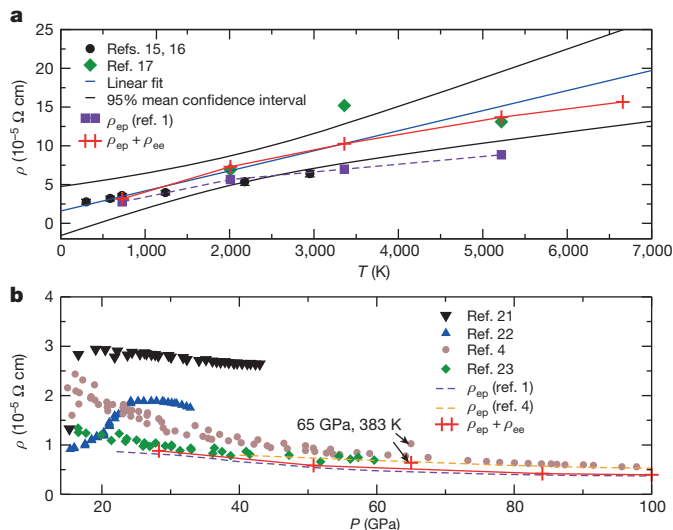


Figure 2 | Our computed resistivities of hcp iron are compared with experimental results. **a**, Resistivity versus temperature along the Hugoniot from shock data^{15–17}, electron-phonon scattering from DFPT calculations¹ (violet), and electron-phonon scattering (DFPT) plus electron-electron scattering (DFT + DMFT) (red). The blue line is the linear fit of shock compression data, and the black lines are the 95% mean confidence interval. **b**, Resistivity versus pressure at $T = 300 \text{ K}$ and at $P = 65 \text{ GPa}$, $T = 383 \text{ K}$. Previous diamond anvil cell experimental results^{4,21–23} are compared with the DFPT calculations of ref. 1 (violet dashed lines) and ref. 4 (orange dashed lines). The data of refs 21, 22 were analysed in ref. 1. Red plus symbols indicate the sum of DFPT electron-phonon¹ and our DFT + DMFT electron-electron resistivities. The statistical 1σ error bars of total resistivities are smaller than their symbols. Values are given in Extended Data Tables 1 and 2.

Direct comparisons of our results with shock compression and diamond anvil cell experimental resistivities are provided in Fig. 2. Along the Hugoniot, we find that our electron-phonon plus electron-electron resistivities are linear at high temperatures (Fig. 2a). So we fitted the shock data along the Hugoniot to a straight line as well, and also derived the 95% mean confidence interval from the data. It is not justified to fit a higher-order function to the data given the experimental scatter. The total resistivity $\rho_{\text{ep}} + \rho_{\text{ee}}$, which is the summation of the density functional perturbation theory (DFPT) and DFT + DMFT results, nicely overlaps the best-fitting line of the shock data. In contrast, the ρ_{ep} line falls below the confidence interval, showing that it is a poor model to explain the results. We note that Keeler *et al.*'s^{15,16} error bars are too small, both from the scatter of their own data, and from Bi *et al.*¹⁷, which provided no error estimates. Furthermore, Bi *et al.* suggested that Keeler *et al.*'s values were systematically low, owing to shunting of the current. Our total resistivities $\rho_{\text{ee}} + \rho_{\text{ep}}$ agree well within experimental error with diamond anvil cell experimental results at room temperature^{4,21–23}, and also at the $P = 65 \text{ GPa}$, $T = 383 \text{ K}$ point of ref. 4. Our results are in slightly better agreement with Seagle *et al.*²³ than with Gomi *et al.*⁴, but this difference probably represents the experimental uncertainty, since both are state-of-the-art experiments. At room temperature the resistivity from the electron-electron scattering is insignificant relative to that from the electron-phonon scattering. We expect resistivity contributions also from defects and grain boundaries, and the DFT electron-phonon values do not include contributions from antiferromagnetic correlations^{24,25}, which are expected to be important at moderate to low temperatures. We find that the temperature dependence of the resistivity is much more important than changes with pressure.

When the mean-free path is comparable to the lattice spacings, saturation in resistivity is expected at the Ioffe–Regel value for the electron-phonon component. We estimate the saturation resistivity to be $11.4 \times 10^{-5} \Omega \text{ cm}$ at the core density, which is higher than our estimated resistivity for electron-phonon scattering at $\rho_{\text{ep}} = 9.15 \times 10^{-5} \Omega \text{ cm}$

at 6,000 K. Since the resistivity from electron–electron scattering may exceed the Ioffe–Regel value²⁶, we do not expect saturation effects to be important at the core conditions.

We estimate the thermal conductivity using the Wiedemann–Franz law ($k = LT/\rho$, with Lorentz parameter $L = 2.44 \times 10^{-8} \text{ W } \Omega \text{ K}^{-2}$), giving about $105 \text{ W m}^{-1} \text{ K}^{-1}$ at temperatures from 4,000 K to 7,000 K. Earth's core is not pure crystalline iron but is liquid and contains light elements of the order of 10% by mass. Since the light elements will decrease the electrical and thermal conductivities, this thermal conductivity is close to previously accepted values¹⁰. Furthermore, resistivity increases with melting²⁷, so that there is now no problem driving the dynamo with thermal convection. Although the absolute values of the core resistivity and thermal conductivity cannot be constrained exactly owing to uncertainties in temperature and composition, it is clear (1) that electron–electron scattering is an important component, and (2) that including electron–electron scattering removes any problem with core conductivity being too high to explain the geodynamo. Thus, the transport crisis is solved.

Contrary to general belief that at high temperatures the resistivity of transition metals comes mainly from electron–phonon scattering⁹, our DFT + DMFT computations have shown that the electron–electron scattering is as important as the electron–phonon scattering in hexagonal close packed (hcp) iron. According to the Fermi-liquid theory, at low temperature T the resistivity of metals from electron–electron scattering $\rho_{\text{ee}}(T)$ is proportional to T^2 . Mott suggested that the T^2 behaviour would have a broad crossover region before saturating, but gave no theory for the form, nor has one yet been developed⁸. Above 2,000 K, we find $\rho_{\text{ee}}(T)$ in Fig. 1 to be linear with temperature at constant volume. Interestingly, the linear T dependence of resistivity is widely observed in correlated materials, including high-temperature superconducting cuprates²⁸, heavy Fermion and other correlated metals^{26,29}. In DMFT simulations of the Hubbard model^{126,29}, linear- T resistivity arises from the linear- T dependence of quasiparticle weight at temperatures above the Fermi-liquid coherent energy scale. However, in hcp iron we find that the quasiparticle weight is only weakly dependent on temperature. As shown in Fig. 3a, the conduction electron scattering rate Γ is linear with T above 2,000 K.

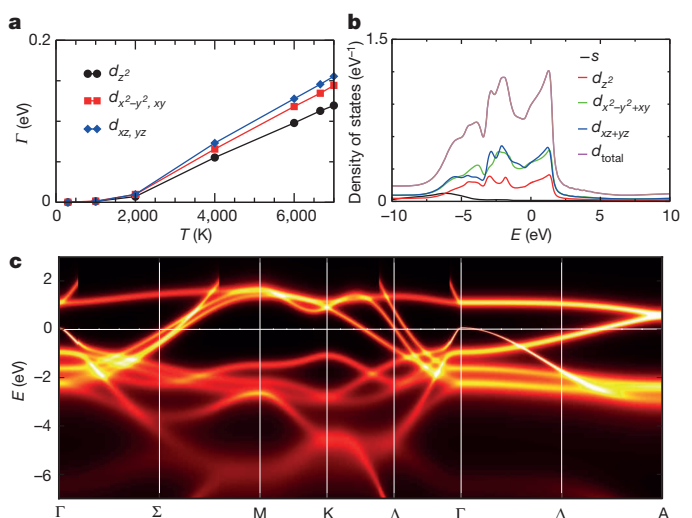


Figure 3 | Scattering rates, density of states and spectral function at Earth's core density of hcp iron. **a**, Orbitaly resolved scattering rates as a function of temperature. $\Gamma_{d_{z^2}}$, $\Gamma_{d_{x^2-y^2, xy}}$ and $\Gamma_{d_{xz, yz}}$ represent the scattering rates on respective d orbitals. **b**, Density of states of s , d_{z^2} , $d_{x^2-y^2, xy}$, $d_{xz, yz}$ and d_{total} orbitals at 6,000 K. **c**, Spectral function $A(k, E)$, where k is the wave vector and E is the electron energy relative to the Fermi level, at 6,000 K. The x axis is the k -path in the first Brillouin zone of the hcp lattice. The statistical 1σ error bars of the scattering rates are smaller than their symbols.

Since the density of s electrons at the Fermi level, $N_s(E_F)$, is very small (Fig. 3b) and in our calculations $N_s(E_F)/[N_s(E_F) + N_d(E_F)] < 1\%$ (where E_F is the Fermi energy and $N_{s,d}$ is the partial density of states) at all temperatures, $\rho_{ee}(T)$ is determined mostly by the scattering rates of d -electrons in hcp iron. We suggest that the linear- T scattering in iron arises from scattering off thermally excited local states which originate from strong electron–electron interactions, a process not included in Fermi-liquid theory. In contrast, the linear resistivity from electron–phonon scattering comes from the near-linear dependence on number of phonons (the quantized lattice vibrations) with temperature.

We find correlated bands in the low-energy region, all being iron $3d$ states (Fig. 3b, 3c). The correlated states at the Fermi level are the origin of large electron–electron scattering and substantial electron–electron resistivity. We expect some other transition metals to have incoherent states at E_F and to show similar behaviour, and those with sharp quasi-particle states at E_F to have normal behaviour with dominant electron–phonon resistivity.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 April; accepted 11 November 2014.

- Sha, X. & Cohen, R. E. First-principles studies of electrical resistivity of iron under pressure. *J. Phys. Condens. Matter* **23**, 075401 (2011).
- de Koker, N., Steinle-Neumann, G. & Vlček, V. Electrical resistivity and thermal conductivity of liquid Fe alloys at high P and T, and heat flux in Earth's core. *Proc. Natl Acad. Sci. USA* **109**, 4070–4073 (2012).
- Pozzo, M., Davies, C., Gubbins, D. & Alfè, D. Thermal and electrical conductivity of iron at Earth's core conditions. *Nature* **485**, 355–358 (2012).
- Gomi, H. *et al.* The high conductivity of iron and thermal evolution of the Earth's core. *Phys. Earth Planet. Inter.* **224**, 88–103 (2013).
- Buffett, B. Geomagnetism under scrutiny. *Nature* **485**, 319–320 (2012).
- Olson, P. The new core paradox. *Science* **342**, 431–432 (2013).
- Haule, K., Yee, C. & Kim, K. Dynamical mean-field theory within the full-potential methods: electronic structure of CeIrIn_5 , CeCoIn_5 , and CeRhIn_5 . *Phys. Rev. B* **81**, 195107 (2010).
- Mott, N. F. Electrons in transition metals. *Adv. Phys.* **13**, 325–422 (1964).
- Allen, P. B. Electron transport. *Contemp. Concepts Condens. Matter Sci.* **2**, 165–218 (2006).
- Stacey, F. D. & Anderson, O. L. Electrical and thermal conductivities of Fe–Ni–S alloy under core conditions. *Phys. Earth Planet. Inter.* **124**, 153–162 (2001).
- Pozzo, M., Davies, C., Gubbins, D. & Alfè, D. Thermal and electrical conductivity of solid iron and iron–silicon mixtures at Earth's core conditions. *Earth Planet. Sci. Lett.* **393**, 159–164 (2014).
- Elsasser, W. M. Induction effects in terrestrial magnetism. Part II. The secular variation. *Phys. Rev.* **70**, 202–212 (1946).
- Yin, Z. P., Haule, K. & Kotliar, G. Kinetic frustration and the nature of the magnetic and paramagnetic states in iron pnictides and iron chalcogenides. *Nature Mater.* **10**, 932–935 (2011).
- Mandal, S., Cohen, R. E. & Haule, K. Strong pressure dependent electron–phonon coupling in FeSe. *Phys. Rev. B* **89**, 220502 (2014).
- Keeler, R. N. & Mitchell, A. C. Electrical conductivity, demagnetization, and the high-pressure phase transition in shock-compressed iron. *Solid State Commun.* **7**, 271–274 (1969).
- Keeler, R. N. & Royce, E. B. in *Physics of High Energy Density* (eds Caldirola, P. & Knoepfel, H.) 106–125 (Proceedings of the International School of Physics Enrico Fermi Vol. 48, 1971).
- Bi, Y., Tan, H. & Jing, F. Electrical conductivity of iron under shock compression up to 200 GPa. *J. Phys. Condens. Matter* **14**, 10849–10854 (2002).
- Lay, T., Hernlund, J. & Buffett, B. A. Core–mantle boundary heat flow. *Nature Geosci.* **1**, 25–32 (2008).
- Wu, B., Driscoll, P. & Olson, P. A statistical boundary layer model for the mantle D'' region. *J. Geophys. Res.* **116**, B12112 (2011).
- Tarduno, J. A. *et al.* Geodynamo, solar wind, and magnetopause 3.4 to 3.45 billion years ago. *Science* **327**, 1238–1240 (2010).
- Balchan, A. S. & Drickamer, H. G. High pressure electrical resistance cell, and calibration points above 100 kilobars. *Rev. Sci. Instrum.* **32**, 308, <http://dx.doi.org/10.1063/1.1717350> (1961).
- Reichlin, R. L. Measuring the electrical resistance of metals to 40 GPa in the diamond–anvil cell. *Rev. Sci. Instrum.* **54**, 1674, <http://dx.doi.org/10.1063/1.1137308> (1983).
- Seagle, C. T., Cottrell, E., Fei, Y., Hummer, D. R. & Prakapenka, V. B. Electrical and thermal transport properties of iron and iron–silicon alloy at high pressure. *Geophys. Res. Lett.* **40**, 5377–5381 (2013).
- Steinle-Neumann, G., Cohen, R. E. & Stixrude, L. Magnetism in iron as a function of pressure. *J. Phys. Condens. Matter* **16**, S1109–S1119 (2004).
- Steinle-Neumann, G., Stixrude, L. & Cohen, R. E. Magnetism in dense hexagonal iron. *Proc. Natl Acad. Sci. USA* **101**, 33–36 (2004).
- Deng, X. *et al.* How bad metals turn good: spectroscopic signatures of resilient quasiparticles. *Phys. Rev. Lett.* **110**, 086401 (2013).
- Deng, L., Seagle, C., Fei, Y. & Shahar, A. High pressure and temperature electrical resistivity of iron and implications for planetary cores. *Geophys. Res. Lett.* **40**, 33–37 (2013).
- Gurvitch, M. & Fiory, A. T. Resistivity of $\text{La}_{1.825}\text{Sr}_{0.175}\text{CuO}_4$ and $\text{YBa}_2\text{Cu}_3\text{O}_7$ to 1100 K: absence of saturation and its implications. *Phys. Rev. Lett.* **59**, 1337–1340 (1987).
- Xu, W., Haule, K. & Kotliar, G. Hidden Fermi liquid, scattering rate saturation, and Nernst effect: a dynamical mean-field theory perspective. *Phys. Rev. Lett.* **111**, 036401 (2013).
- Stixrude, L., Wasserman, E. & Cohen, R. E. Composition and temperature of Earth's inner core. *J. Geophys. Res.* **102**, 24729–24739 (1997).
- Sha, X. & Cohen, R. E. First-principles thermal equation of state and thermoelasticity of hcp Fe at high pressures. *Phys. Rev. B* **81**, 094105 (2010).
- Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).

Acknowledgements This work is supported by National Science Foundation (NSF) grants EAR-1214807, DMS-1025392, and DMR-1405303. R.E.C. is supported by the Carnegie Institution and the European Research Council Advanced Grant ToMCoT. K.H. is supported by NSF grant DMR-1405303. This research used the NSF Extreme Science and Engineering Discovery Environment (XSEDE) supercomputer 'Stampede', and also used the resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the US Department of Energy under contract number DE-AC05-00OR22725. R.E.C. and P.Z. thank I. Mazin, S. Labrosse and R. Caracas for discussions. We also acknowledge J. Robb for assistance with the manuscript preparation.

Author Contributions R.E.C. designed the project, P.Z. performed the computations, P.Z., R.E.C. and K.H. analysed the results and prepared the paper. The DFT + DMFT code was developed by K.H.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.E.C. (rcohen@carnegiescience.edu).

METHODS

Code availability. The DFT + DMFT code was developed by K.H. and is available at <http://hauleweb.rutgers.edu/downloads/>.

The DFT + DMFT formalism. In the DFT + DMFT method^{33,34} a functional (equation 118 of ref. 33) that includes all local two particle irreducible skeleton diagrams is optimized. The interaction Hamiltonian is given by the Slater form (equation 28 of ref. 7) with the Slater integral $F^0 = U$, $F^2 = (14/1.625)J$, and $F^4 = (8.75/1.625)J$, where U is the Hubbard parameter and J is the Hund's coupling. The double counting energy E_{dc} is calculated from the fully localized limit³⁵ formula $E_{dc} = U \left(n_{cor}^0 - \frac{1}{2} \right) - \frac{J}{2} (n_{cor}^0 - 1)$ and n_{cor}^0 is the nominal electron occupancy of the correlated atom (iron in this case). We also tested the around-mean-field double-counting³⁶ and found an increment of resistivity of up to 16% at core conditions, which is not significant. Our DFT calculations show $n_{cor} = 6.6$ at the core density of iron, so we choose $n_{cor}^0 = \text{Int}(n_{cor}) = 7$, where Int means choosing the nearest integer number of n_{cor} . We tested $n_{cor}^0 = 6$ and 8, but did not find large changes, and $n_{cor}^0 = 7$ gives the lowest resistivity. More details are given in ref. 7. In our calculation, we choose an energy window of ± 10 eV around the Fermi level E_F for the projector. We use the continuous time quantum Monte Carlo method to sample all diagrams in the hybridization expansion, as described in detail in refs 37 and 38.

The all-electron LAPW (linearized augmented plane wave) WIEN2K code³⁹ is used for the DFT calculations, with the Wu–Cohen exchange–correlation potential⁴⁰. A test at core conditions with the Perdew–Burke–Ernzerhof exchange–correlation functional did not change the resistivity much (only by 2.5%), indicating that our results are robust. The k -space summation is on a $12 \times 12 \times 12$ grid using a modified tetrahedron integration scheme⁴¹. The cut-off energy separating the core from valence states is -9.0 Ry. $R_{mt}K_{max}$ is 9.0 (where R_{mt} is the smallest atomic sphere radius and K_{max} is the maximum number of wavevectors used) and the magnitude of the largest vector GMAX is 19.0. There is no spin-orbital coupling/splitting in our calculation.

The DFT + DMFT method iterates as follows: (1) the lattice problem is solved as in DFT, but with added self-energy for the correlated states (which is zero for the first iteration), which makes the problem non-Hermitian and frequency dependent. The eigenvalues, wavefunctions, charge density and potential are output. Step 1 can be iterated as an inner loop. (2) The impurity levels E_{imp} and the hybridization function $\Delta(\omega)$ between the lattice and the impurity are computed and input into the continuous time quantum Monte Carlo impurity solver to find the DMFT solution. Step 2 can be iterated as an inner loop. (3) The self-energy and the electron density are updated; the new self-energy and electron density are inserted into step (1) and (2) for the next DFT + DMFT iteration.

For Earth's core density of hcp iron we use 13.04 g cm^{-3} and the corresponding lattice volume of 47.8 atomic units^{42,43}. The pressure–temperature relationship at this density is presented in Extended Data Fig. 1. The lattice parameter ratio c/a is 1.615 at all volumes⁴⁴. We used a Hubbard $U = 5$ eV and Hund's coupling $J = 0.943$ eV on the basis of numerous previous studies of iron compounds. We checked that $U = 2$ eV gives very similar results; the resistivity is weakly dependent on U (the largest difference is smaller than 12%). Reducing J to zero halves the computed electron–electron resistivity, showing that the Hund's coupling is quantitatively important but not solely responsible for the scattering. Reducing U and J to near zero reduces the electronic resistivity to near zero, as expected. Without the electron–electron correlations, the electron–phonon interactions would dominate, as previously believed.

Accuracy of the DFT + DMFT method. The DFT + DMFT method combines the accurate treatment of the many-body physics as well as a fully self-consistent treatment of the crystal and atomic bonding and hybridization. The DFT + DMFT method is one of the most important advances in numerical simulation of condensed matter physics. By introducing correlation effects, this method works in the region where DFT fails to predict experimental results. Its accuracy has been proved in research on various correlated electron materials from transition metals and their compounds to heavy fermion materials^{33,45}. At ambient conditions the transition-metal oxide FeO is an insulator but DFT predicts it to be a metal⁴⁶. In contrast, DFT + DMFT not only makes FeO an insulator at ambient conditions, but also successfully predicted the existence of a metallic phase at high pressure⁴⁷. DFT + DMFT has also been used in research on heavy fermion materials. Using DFT + DMFT, Shim *et al.*⁴⁸ identified the ground-state electronic configurations of curium and plutonium. They found that curium has a single-valence ground state with magnetic ordering, whereas plutonium has a ground state that comes from superposition of two atomic valences. The different magnetic properties of curium and plutonium are explained by the interplay between their ground-state electronic configurations, the electronic itinerancy and localization, as well as the spin-orbit coupling. The same group also investigated CeIrIn₅ using DFT + DMFT⁴⁹, where they found the numerically calculated temperature resolved spectral functions to

be in good agreement with experimental results. The consistency between numerical simulations and experimental results enabled them to explain the experimentally observed features in the optical conductivity. The accuracy of DFT + DMFT predictions are not limited to the single-particle level. In an inelastic neutron scattering experiment of the iron pnictide BaFe_{1.9}Ni_{0.1}As₂ by ref. 50, their experimental data were compared with the dynamical magnetic susceptibility from DFT + DMFT calculations. They found systematic consistency between the experimental and the numerical results at different energy and moment slices. From these results, they confirm that magnetic excitations in the iron pnictide BaFe_{1.9}Ni_{0.1}As₂ are partially localized, which indicates the strongly correlated nature of this high-temperature superconducting material.

Scattering rate and analytic continuation. In our DFT + DMFT calculations the scattering rate of d electrons at the Fermi level is given by:

$$\Gamma_{E_F}^{\alpha} = -Z_{K_F}^{\alpha} \text{Im}\Sigma^{\alpha}(\omega)|_{\omega \rightarrow 0} \quad (2)$$

in which $Z_{K_F}^{\alpha} = [1 - (\partial \text{Re}\Sigma^{\alpha}(\omega)/\partial \omega)]^{-1}|_{\omega \rightarrow 0}$, α is the d -orbital index of d_{z^2} , $d_{x^2-y^2}$, xy and $d_{xz,yz}$ and $\Sigma(\omega)$ is the self-energy in real frequency from maximum entropy (MaxEnt) analytic continuation⁵¹ (Extended Data Figs 2 and 3).

Another DFT + DMFT computation for hcp iron (atomic volume of 47.6 atomic units, hcp lattice ratio $c/a = 1.6$, $U = 3.37$ eV, $J = 0.93$ eV, projection energy window $[-10.8$ eV, 4.0 eV] around the Fermi level, around-mean-field double counting) is presented in ref. 52. We duplicate all of their results by exactly following their methods. In ref. 52 the scattering rate of hcp iron is calculated by extrapolating the imaginary frequency to zero $\Gamma(T, i0^+) = -Z(T, i\omega_n) \text{Im}[\Sigma(T, i\omega_n)]|_{i\omega_n \rightarrow i0^+}$. They claim that hcp iron is in the Fermi-liquid state up to 5,800 K with a quadratic scattering rate in temperature. Although we agree with their results at room temperature, we find very different behaviour at high temperatures. Their extrapolation in imaginary frequency is poorly constrained at high temperatures. At Earth's core temperature ($T = 6,000$ K), the first positive imaginary frequency is at $\omega_0 = 1.62$ eV, and there are only two points that could be used in extrapolation below 5 eV. This makes the self-energy, and consequently the scattering rate, at $i0^+$ depend heavily on the choice of extrapolation. In Extended Data Fig. 2, three methods (linear, cubic and Akima) are used to extrapolate the imaginary part of the self-energy to $i0^+$. The results from the three extrapolations are distributed over a wide range. In Extended Data Fig. 2, the imaginary part of self-energy at $i0^+$ ranges from -0.23 eV to -0.12 eV, giving 100% uncertainty for the absolute values. Such uncertainty at high temperatures leads ref. 52 to conclude that hcp iron is a Fermi liquid at high temperatures, which is contrary to our results, owing to insufficient accuracy in their analysis. In contrast, the self-energy from our MaxEnt has a very dense mesh, where the smallest energy scale is 0.0025 eV, as shown in the inset of Extended Data Fig. 3. There is a $\omega = 0$ point, so extrapolation is no longer needed.

Three independent methods were used to check the stability and accuracy of our analytic continuation; we used both the Padé and the singular value decomposition methods in addition to MaxEnt. As presented in Extended Data Fig. 3 and its inset, the imaginary part of the self-energies from the three analytic continuations agree in the low-energy region needed for the conductivity. At energies around the Fermi level, the three analytic continuations give identical self-energies. This proves that our MaxEnt results are precise and stable.

Optical conductivity calculation in DMFT. The optical conductivity is calculated using the self-energy on the real frequency axis from the MaxEnt, with the low frequency limit giving the direct-current conductivity. In DMFT, since the vertex corrections to conductivity can be safely omitted⁵³, the formula we use in the optical conductivity calculation is⁴⁵

$$\text{Re}\sigma_{\mu,\nu}(\omega) = \pi e^2 \sum_k \int_{-\infty}^{\infty} d\epsilon \frac{f(\epsilon - \omega) - f(\epsilon)}{\omega} \text{Tr}[\rho_k(\epsilon) V_{k,\mu}^{\dagger} \rho_k(\epsilon - \omega) V_{k,\nu}] \quad (3)$$

where μ, ν are direction indices, $\rho_k(\epsilon) = \frac{G_k^{\dagger}(\epsilon) - G_k(\epsilon)}{2\pi i}$, the velocity vector $V_{ij}^{k,\mu} = -\frac{i}{m} \langle \psi_{k,i} | \nabla_{\mu} | \psi_{k,j} \rangle$ and i, j are orbital indices, the fermionic distribution function $f(\epsilon) = (e^{\beta(\epsilon - \mu)} + 1)^{-1}$ and the trace is over all valence states.

Equation of state. Our resistivity results at core conditions are independent of any equation of state. We use the known density⁴³ and the temperature of Earth's core. For our comparisons with experiments under other conditions, we estimate the pressures of our DFT + DMFT calculations from the thermal equation of state given by ref. 31. The Hugoniot line is from the same paper³¹. The pressure–temperature relationship at Earth's core density and along the Hugoniot line are shown in Extended Data Fig. 1.

Extrapolations. In Fig. 1 we estimated the resistivities of iron at Earth's core density using the systematics of ref. 10 as well as three sets of parameters (P, T_m) (where T_m is the melting temperature of iron) along iron's melting curve^{54–56}. Stacey and

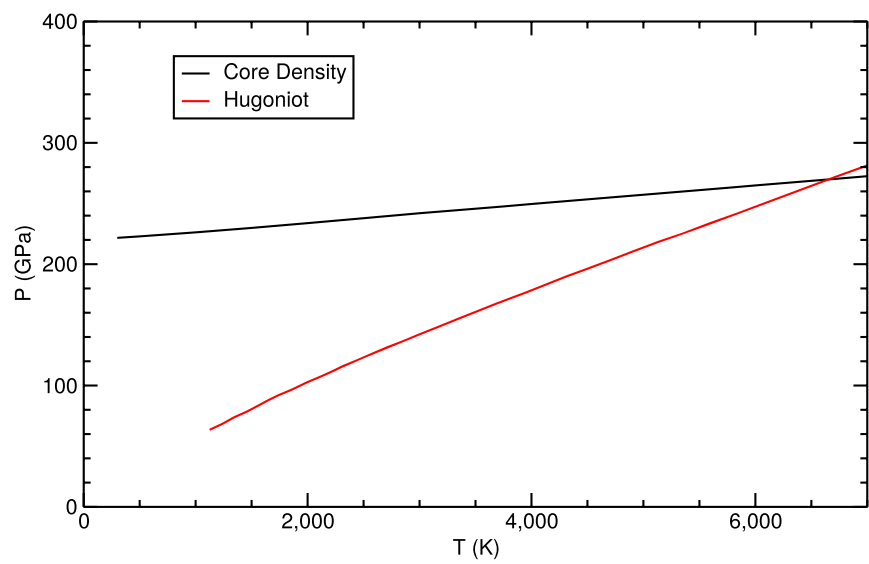
Anderson¹⁰ assume that the resistivity of iron (1) is constant along the melting line at $13.5 \times 10^{-5} \Omega \text{ cm}$ and (2) is a linear function of temperature at constant pressure. In Extended Data Table 1a the temperature T_c at the core density corresponding to (P, T_m) is derived from the equation of state in ref. 31 (see Extended Data Fig. 1). The resistivities at the core density are given by $\rho_c = \rho_m T_c / T_m$, where ρ_m is the melting resistivity. We also tested the effectiveness of Stacey and Anderson's systematics¹⁰ using our resistivity data at the atomic volume of 45 atomic units and temperature 6,000 K. The derived melting resistivity is always close to $13.5 \times 10^{-5} \Omega \text{ cm}$. Interestingly, our calculations support the systematics of Stacey and Anderson¹⁰, in spite of the importance we find of electron–electron scattering, and their assumption that electron–phonon scattering would give scaling with the melting curve.

The shock compression experiment extrapolation point in Fig. 1 is derived from the formula:

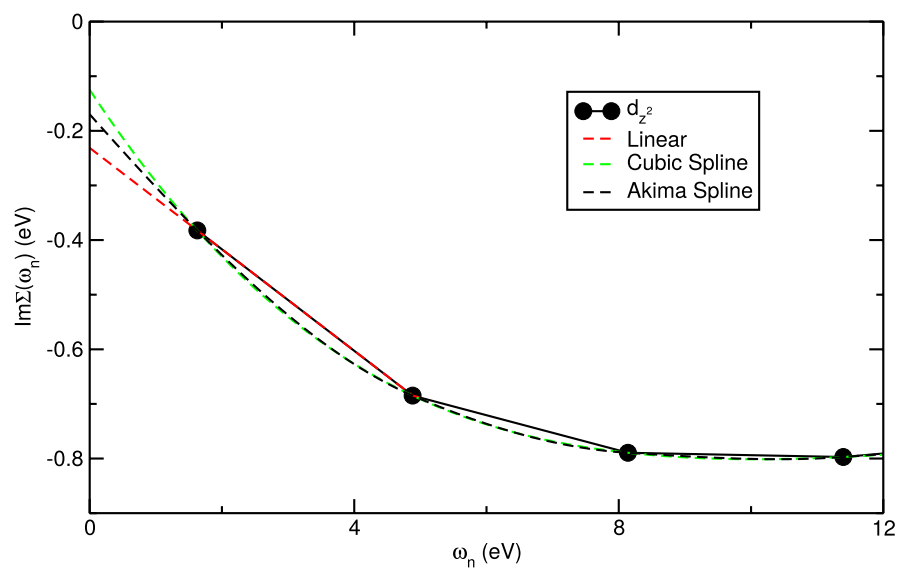
$$\rho(T) = 1.58 + 2.59 \times 10^{-3} T \quad (4)$$

which is the best linear fit of resistivity data from previous shock experiments by Keeler *et al.*^{15,16} and Bi *et al.*¹⁷. $\rho(T)$ is in units of $10^{-5} \Omega \text{ cm}$. The pressure and temperature of this point come from the Hugoniot line at Earth's core density, as given in Extended Data Fig. 1.

33. Kotliar, G. *et al.* Electronic structure calculations with dynamical mean-field theory. *Rev. Mod. Phys.* **78**, 865 (2006).
34. Georges, A., Kotliar, G., Krauth, W. & Rozenberg, M. J. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Mod. Phys.* **68**, 13 (1996).
35. Anisimov, V. I., Aryasetiawan, F. & Lichtenstein, A. I. First-principles calculations of the electronic structure and spectra of strongly correlated systems: the LDA+U method. *J. Phys. Condens. Matter* **9**, 767–808 (1997).
36. Czyżyk, M. T. & Sawatzky, G. A. Local-density functional and on-site correlations: The electronic structure of La_2CuO_4 and LaCuO_3 . *Phys. Rev. B* **49**, 14211 (1994).
37. Werner, P., Comanac, A., de'Medici, L., Troyer, M. & Millis, A. J. Continuous-time solver for quantum impurity models. *Phys. Rev. Lett.* **97**, 076405 (2006).
38. Haule, K. Quantum Monte Carlo impurity solver for cluster dynamical mean-field theory and electronic structure calculations with adjustable cluster base. *Phys. Rev. B* **75**, 155113 (2007).
39. Blaha, P., Schwarz, K., Madsen, G. K. H., Kvasnicka, K. & Luitz, J. in *Wien2K* (ed. Schwarz, K.) (Technische Universität Wien, 2001).
40. Wu, Z. & Cohen, R. E. More accurate generalized gradient approximation for solids. *Phys. Rev. B* **73**, 235116 (2006).
41. Blöchl, P. E., Jepsen, O. & Andersen, O. K. Improved tetrahedron method for Brillouin-zone integrations. *Phys. Rev. B* **49**, 16223 (1994).
42. Sha, X. & Cohen, R. E. Elastic isotropy of ϵ -Fe under Earth's core conditions. *Geophys. Res. Lett.* **37**, L10302 (2010).
43. Dziewonski, A. M. & Anderson, D. L. Preliminary reference Earth model. *Phys. Earth Planet. Inter.* **25**, 297–356 (1981).
44. Sha, X. & Cohen, R. E. Thermal effects on lattice strain in ϵ -Fe under pressure. *Phys. Rev. B* **74**, 064103 (2006).
45. Basov, D. N., Averitt, R. D., van der Marel, D., Dressel, M. & Haule, K. Electrodynamics of correlated electron materials. *Rev. Mod. Phys.* **83**, 471–541 (2011).
46. Cohen, R. E., Fei, Y., Downs, R., Mazin, I. I. & Isaak, D. G. in *High-Pressure Materials Research Materials Research Society Proceedings* (eds Wentzcovitch, R., Hemley, R. J., Nellis, W. J. & Yu, P.) Vol. 499, 27–37 (Materials Research Society, 1998).
47. Ohta, K. *et al.* Experimental and theoretical evidence for pressure-induced metallization in FeO with the rock-salt type structure. *Phys. Rev. Lett.* **108**, 026403 (2012).
48. Shim, J. H., Haule, K. & Kotliar, G. Fluctuating valence in a correlated solid and the anomalous properties of δ -plutonium. *Nature* **446**, 513–516 (2007).
49. Shim, J. H., Haule, K. & Kotliar, G. Modelling the localized to itinerant electronic transition in the heavy fermion system CeIrIn_5 . *Science* **318**, 1615–1617 (2007).
50. Liu, M. *et al.* Nature of magnetic excitations in superconducting $\text{BaFe}_{1.9}\text{Ni}_{0.1}\text{As}_2$. *Nature Phys.* **8**, 376 (2012).
51. Jarrell, M. & Gubernatis, J. E. Bayesian inference and the analytic continuation of imaginary time quantum Monte Carlo data. *Phys. Rep.* **269**, 133–195 (1996).
52. Pourvorskii, L. V. *et al.* Electronic properties and magnetism of iron at the Earth's inner core conditions. *Phys. Rev. B* **87**, 115130 (2013).
53. Khurana, A. Electrical conductivity in the infinite-dimensional Hubbard model. *Phys. Rev. Lett.* **64**, 1990 (1990).
54. Yoo, C. S., Akella, J., Campbell, A. J., Mao, H. K. & Hemley, R. J. Phase diagram of iron by in situ X-ray diffraction: implications for Earth's core. *Science* **270**, 1473–1475 (1995).
55. Brown, J. M. & McQueen, R. G. Phase transitions, Grüneisen parameter, and elasticity for shocked iron between 77 GPa and 400 GPa. *J. Geophys. Res.* **91**, 7485–7494 (1986).
56. Nguyen, J. H. & Holmes, N. C. Melting of iron at the physical conditions of the Earth's core. *Nature* **427**, 339–342 (2004).

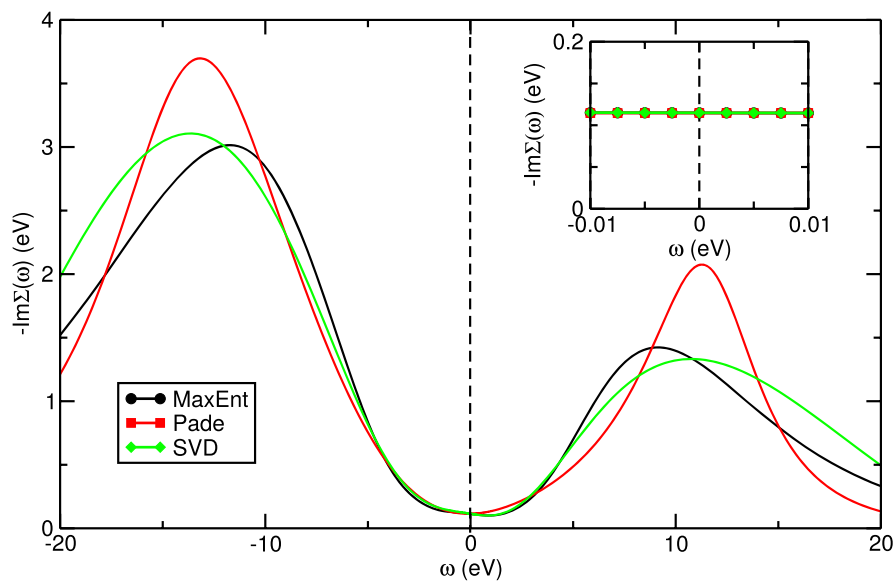


Extended Data Figure 1 | Pressure versus temperature relationship of hcp iron at Earth’s core density and along the Hugoniot line³¹. The two lines cross at $P = 269.9$ GPa, $T = 6,658$ K at Earth’s core density.



Extended Data Figure 2 | Extrapolation of $\text{Im}\Sigma(i\omega_n)$ to zero imaginary frequency. The self-energy is from the d_{z^2} orbital of hcp iron at Earth's core density and 6,000 K. Three extrapolation methods are used: linear, cubic spline

and Akima spline. The imaginary part of self-energy at $i0^+$ ranges from -0.23 eV to -0.12 eV.



Extended Data Figure 3 | The imaginary part of self-energies in real frequency on the d_{z^2} orbital of hcp iron at Earth's core density and 6,000 K. The self-energies are from three analytic continuation methods: MaxEnt, Padé and singular value decomposition. The inset shows the same imaginary

part of self-energies in energy range $[-0.01 \text{ eV}, 0.01 \text{ eV}]$ around the Fermi level. The self-energies from three analytic continuation methods agree at the low-energy region.

Extended Data Table 1 | Resistivities from extrapolations and previous experiments

a	P (GPa)	T_m (K)	T_c (K)	ρ ($10^{-5} \Omega \text{ cm}$)	
	235	5495	2148	5.28	
	243	5572	3141	7.61	
	260	5737	5361	12.62	
b	V ($\text{bohr}^3/\text{atom}$)	P (GPa)	T (K)	ρ ($10^{-5} \Omega \text{ cm}$)	\pm ($10^{-5} \Omega \text{ cm}$)
	47.8	269.9	6658	18.83	3.07
c	Source	V ($\text{bohr}^3/\text{atom}$)	P (GPa)	T (K)	ρ ($10^{-5} \Omega \text{ cm}$)
	de Koker et al.	47.8	264.9	6000	6.56
	Pozzo et al.	47.8	264	5865	6.65
d	V ($\text{bohr}^3/\text{atom}$)	P (GPa)	T (K)	ρ ($10^{-5} \Omega \text{ cm}$)	\pm ($10^{-5} \Omega \text{ cm}$)
	74.2	17	303	2.78	0.31
	67.0	37	585	3.19	0.31
	63.0	44.4	728	3.59	0.24
	61.7	64.3	1240	3.95	0.28
	57.2	110	2180	5.35	0.45
	55.4	140	2950	6.41	0.42
e	V ($\text{bohr}^3/\text{atom}$)	P (GPa)	T (K)	ρ ($10^{-5} \Omega \text{ cm}$)	
	57.9	101.1	2010	6.9	
	55.1	146.7	3360	15.2	
	52.3	208.0	5220	13.1	

a, The extrapolated resistivities in Fig. 1 at Earth's core density using the systematics of ref. 10. **b**, The extrapolated resistivity in Fig. 1 at Earth's core density on the Hugoniot. **c**, The resistivities from DFT + MD calculations in Fig. 1 at Earth's core density, extracted from refs 2 and 3. **d**, The atomic volumes, pressures, temperatures and resistivities from shock compression experiments^{15,16} in Fig. 2a. **e**, The atomic volumes, pressures, temperatures and resistivities from shock compression experiments by ref. 17 in Fig. 2a.

Extended Data Table 2 | The atomic volumes, pressures, temperatures and resistivities from our study in Fig. 1 and Fig. 2

	V ($\text{bohr}^3/\text{atom}$)	P (GPa)	T (K)	ρ_{ep} ($10^{-5} \Omega \text{ cm}$)	ρ_{ee} ($10^{-5} \Omega \text{ cm}$)	$\pm\rho_{ee}$ ($10^{-6} \Omega \text{ cm}$)	$\rho_{ep} + \rho_{ee}$ ($10^{-5} \Omega \text{ cm}$)
a	47.8	221.7	300	0.30	0.012	0.0025	0.312
	47.8	226.3	1000	1.39	0.14	0.028	1.53
	47.8	233.9	2000	2.94	0.55	0.10	3.49
	47.8	249.6	4000	6.05	2.68	0.48	8.73
	47.8	264.9	6000	9.15	4.75	0.81	13.90
	47.8	269.9	6658	10.18	5.48	1.10	15.66
	47.8	272.6	7000	10.71	5.87	1.17	16.58
	V ($\text{bohr}^3/\text{atom}$)	P (GPa)	T (K)	ρ_{ep} ($10^{-5} \Omega \text{ cm}$)	ρ_{ee} ($10^{-5} \Omega \text{ cm}$)	$\pm\rho_{ee}$ ($10^{-6} \Omega \text{ cm}$)	$\rho_{ep} + \rho_{ee}$ ($10^{-5} \Omega \text{ cm}$)
b	63	44.4	728	2.77	0.38	0.057	3.15
	57.9	101	2010	5.65	1.65	0.31	7.30
	55.1	146.7	3360	6.98	3.28	0.66	10.26
	52.3	208	5220	8.84	4.85	0.87	13.69
	V ($\text{bohr}^3/\text{atom}$)	P (GPa)	T (K)	ρ_{ep} ($10^{-5} \Omega \text{ cm}$)	ρ_{ee} ($10^{-5} \Omega \text{ cm}$)	$\pm\rho_{ee}$ ($10^{-6} \Omega \text{ cm}$)	$\rho_{ep} + \rho_{ee}$ ($10^{-5} \Omega \text{ cm}$)
c	63	28.3	300	0.82	0.063	0.011	0.883
	60	50.8	300	0.54	0.042	0.0076	0.582
	56.5	84.1	300	0.39	0.024	0.0043	0.414
	55	100.0	300	0.37	0.021	0.0040	0.391
	58	65	383	0.59	0.054	0.011	0.644

a, DFT+DMFT calculated resistivities in Fig. 1 at Earth's core density. **b**, DFT + DMFT calculated resistivities in Fig. 2a, along the Hugoniot line. **c**, DFT + DMFT calculated resistivities in Fig. 2b, compared with DAC experimental results. ρ_{ep} is the DFPT-calculated resistivity by ref. 1. ρ_{ee} is the resistivity from our DFT + DMFT study.

Resolving the complexity of the human genome using single-molecule sequencing

Mark J. P. Chaisson¹, John Huddleston^{1,2}, Megan Y. Dennis¹, Peter H. Sudmant¹, Maika Malig¹, Fereydoon Hormozdiari¹, Francesca Antonacci³, Urvashi Surti⁴, Richard Sandstrom¹, Matthew Boitano⁵, Jane M. Landolin⁵, John A. Stamatoyannopoulos¹, Michael W. Hunkapiller⁵, Jonas Korlach⁵ & Evan E. Eichler^{1,2}

The human genome is arguably the most complete mammalian reference assembly^{1–3}, yet more than 160 euchromatic gaps remain^{4–6} and aspects of its structural variation remain poorly understood ten years after its completion^{7–9}. To identify missing sequence and genetic variation, here we sequence and analyse a haploid human genome (CHM1) using single-molecule, real-time DNA sequencing¹⁰. We close or extend 55% of the remaining interstitial gaps in the human GRCh37 reference genome—78% of which carried long runs of degenerate short tandem repeats, often several kilobases in length, embedded within (G+C)-rich genomic regions. We resolve the complete sequence of 26,079 euchromatic structural variants at the base-pair level, including inversions, complex insertions and long tracts of tandem repeats. Most have not been previously reported, with the greatest increases in sensitivity occurring for events less than 5 kilobases in size. Compared to the human reference, we find a significant insertional bias (3:1) in regions corresponding to complex insertions and long short tandem repeats. Our results suggest a greater complexity of the human genome in the form of variation of longer and more complex repetitive DNA that can now be largely resolved with the application of this longer-read sequencing technology.

Data generated by single-molecule, real-time (SMRT) sequencing technology differ drastically from most sequencing platforms because native DNA is sequenced without cloning or amplification, and read lengths typically exceed 5 kilobases (kb). Despite overall lower individual read accuracy (~85%), longer read length facilitates high confidence mapping across a greater percentage of the genome^{11,12}. We generated ~40-fold sequence coverage from a human CHM1 hydatidiform mole using long-read SMRT sequence technology (average mapped read length = 5.8 kb; Supplementary Table 1). We selected a complete hydatidiform mole to sequence because it is haploid, lacking allelic variation, and provides higher effective sequence coverage. We aligned 93.8% of all sequence reads to the human reference genome (GRCh37) using a modified version of BLASR¹¹ (Supplementary Information) and generated local assemblies of the mapped reads using Celera¹³ and Quiver¹⁴, the latter of which leverages estimates of insertion, deletion and substitution probabilities to determine consensus sequences accurately. We compared the consensus sequences of regions with previously sequenced and assembled large-insert bacterial artificial chromosome (BAC) clones generated from CHM1tert (ref. 15). The comparison shows a consensus sequencing concordance of >99.97% (phred quality = 37.5), with 72% of the errors confined to indels within homopolymer stretches (Supplementary Table 3).

We initially assessed whether the mapped reads could facilitate closure of any of the 164 interstitial euchromatic gaps within the human reference genome (GRCh37). We extended into gap regions using a reiterative map-and-assemble strategy, in which SMRT whole-genome sequencing (WGS) reads mapping to each edge of a gap were assembled into a new high-quality consensus, which, in turn, served as a template

for recruiting additional sequence reads for assembly (Supplementary Information). Using this approach, we closed 50 gaps and extended into 40 others (60 boundaries), adding 398 kb and 721 kb of novel sequence to the genome, respectively (Supplementary Table 4). The closed gaps in the human genome were enriched for simple repeats, long tandem repeats, and high (G+C) content (Fig. 1) but also included novel exons (Supplementary Table 20) and putative regulatory sequences based on DNase I hypersensitivity and chromatin immunoprecipitation followed by high-throughput DNA sequencing (ChIP-seq) analysis (Supplementary Information). We identified a significant 15-fold enrichment of short tandem repeats (STRs) when compared to a random sample ($P < 0.00001$) (Fig. 1a). A total of 78% (39 out of 50) of the closed gap sequences were composed of 10% or more of STRs. The STRs were frequently embedded in longer, more complex, tandem arrays of degenerate repeats reaching up to 8,000 bp in length (Extended Data Fig. 1a–c), some of which bore resemblance to sequences known to be toxic to *Escherichia coli*¹⁶. Because most human reference sequences^{17,18} have been derived from clones propagated in *E. coli*, it is perhaps not surprising that the application of a long-read sequence technology to uncloned DNA would resolve such gaps. Moreover, the length and complex degeneracy of these STRs embedded within (G+C)-rich DNA probably thwarted efforts to follow up most of these by PCR amplification and sequencing.

Next, we developed a computational pipeline (Extended Data Fig. 2) to characterize structural variation systematically (structural variation defined here as differences ≥ 50 bp in length, including deletions, duplications, insertions and inversions⁷). Structural variants were discovered by mapping SMRT sequencing reads to the human reference genome¹¹

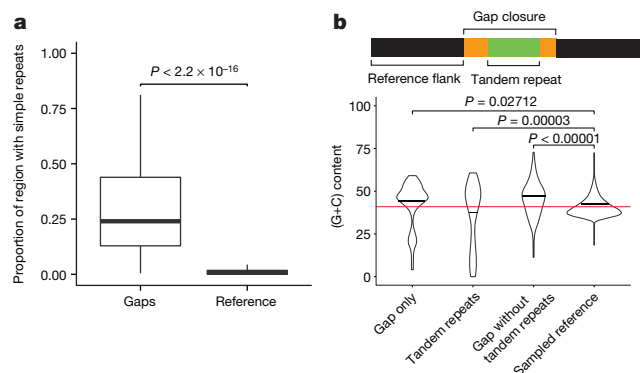


Figure 1 | Sequence content of gap closures. **a**, Gap closures are enriched for simple repeats compared to equivalently sized regions randomly sampled from GRCh37. **b**, Human genome gaps typically consist of (G+C)-rich sequence (yellow) flanking complex (A+T)-rich STRs (green) (empirical P value; Supplementary Information). Red line indicates genomic (G+C) content.

¹Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington 98195, USA. ²Howard Hughes Medical Institute, University of Washington, Seattle, Washington 98195, USA. ³Dipartimento di Biologia, Università degli Studi di Bari 'Aldo Moro', Bari 70125, Italy. ⁴Department of Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA. ⁵Pacific Biosciences of California, Inc., Menlo Park, California 94025, USA.

Table 1 | Structural variation between CHM1 and GRCh37

	Insertion			Deletion			Ins/del	
	Number	Mean length	Total bases	Number	Mean length	Total bases	Total events	Total bases
STR >10 bp	6,007	295	1,771,948	2,986	90	268,075	2.01	6.61
STR ≥ 50 bp	4,289	398	1,706,524	1,530	139	212,957	2.80	8.01
STR >10, < 50 bp	1,718	38	65,424	1,456	38	5,518	1.18	11.86
Tandem repeat	2,760	303	836,474	2,398	182	4,361,598	1.15	0.19
MEI	2,149	497	1,200,647	2,084	428	841,617	1.03	1.43
AluY	859	302	259,810	859	302	259,220	1.00	1.00
LINE/L1Hs	145	2,412	349,780	141	2,411	339,971	1.03	1.03
SVA	457	369	168,762	382	274	104,589	1.20	1.61
HERV	58	338	19,619	60	180	10,779	0.97	1.82
Alu+STR/Alu+mosaic	287	413	118,486	186	262	46,905	1.54	2.53
Inactive	343	226	77,602	456	176	80,153	0.75	0.97
Centromeric satellites	669	693	463,687	817	722	590,223	0.82	0.79
HSAT	46	861	39,604	48	790	37,935	0.96	1.04
ALR	622	681	423,453	769	718	552,288	0.81	0.77
Other	168	112	18,790	277	98	27,144	0.61	0.69
Complex	1,115	1,927	2,148,642	317	2,066	654,834	3.52	3.28
Unannotated	2,386	60	143,598	2,313	62	143,559	1.03	1.00
Total	17,851	398	7,112,381	11,819	271	3,208,633	1.51	2.22
Euchromatic subtotal	15,776	390	6,149,335	10,303	248	2,559,644	1.53	2.40
Euchromatic subtotal (≥50 bp)	9,638	542	5,237,445	6,111	358	2,189,837	1.58	2.39

The statistics of insertion and deletion events in CHM1 compared to GRCh37 are listed by sequence category. Low complexity sequence is divided between STRs and variable number tandem repeats (Supplementary Information). AluY, L1Hs, SVA and HERV are active mobile elements. Alu indel events in conjunction with STR sequences or mosaic Alu are considered separately from solitary AluY mobile element insertions (MEIs). Inactive mobile element insertions include L1P and AluS. Rarely observed elements (<10) are combined as 'Other'. Classes of structural variation showing an insertional bias (>2.5-fold excess in CHM1) are in bold.

and searching for specific mapping signatures (Supplementary Information). At every variant locus, we recruited all uniquely mapping reads, created a local *de novo* assembly, defined breakpoints compared to the human reference, and classified each structural variant by type and probable mechanism (Table 1). We identified a total of 26,079 insertions/deletions ≥50 bp within the euchromatic portion of the genome. Almost all insertion and deletion breakpoints were resolved at the single-base-pair level, generating one of the most comprehensive catalogues of structural variation (47,238 breakpoint positions). A total of 6,796 of the events map within 3,418 genes with a subset of events (169) corresponding to variation in the spliced transcripts of 140 genes (Supplementary Table 9). From all targeted sequencing experiments combined (Supplementary Information) we estimate an overall validation rate of 97%, of which only a fraction can be detected by application of Illumina next-generation sequencing.

Of all copy number differences found, 85% were novel compared to previous studies of structural variation^{7,8,19}, in large part owing to increased ascertainment of smaller variation (average length 497 bp). The effect was most pronounced for insertions in which 92% of all differences had not been previously reported, in contrast to deletions in which 69% of the events were novel (Fig. 2). When comparing the size distribution of insertions and deletions between the two haplotype references, we found that insertions within CHM1 were longer and more abundant with 5,473 additional insertion events when compared to the human reference (Table 1). This difference contributes to a significant insertional bias of 3.9 megabases (Mb) of additional sequence either missing or expanded when compared to the human reference (Table 1). We find a substantial increase in the amount of long, ≥50 bp STR insertions relative to deletions ($P < 2.2 \times 10^{-16}$), including STRs within genes (Supplementary Table 9). In addition to being 2.80 times more frequent than deletions, the STR insertions ≥50 bp are, on average, 2.87 times longer. This asymmetry becomes more pronounced with increasing STR insertion length (Fig. 2b). The genomic distribution of STR insertions is highly non-random being biased to the last 5 Mb of human chromosomes (Extended Data Fig. 3) correlating with recombination rate²⁰ ($r^2 = 0.21$) and human–chimpanzee divergence ($r^2 = 0.20$). We note that 2,285 of these expanded STRs occur within genes, including 11 within an untranslated region (noting shorter insertions in *FMRI* and *C9orf72*, a common mutated locus for amyotrophic lateral sclerosis; Supplementary Information) and two within the coding sequence of genes (*MUC2* and

SAMD1). A total of 189 genes have an STR expansion >1 kb, representing potential sites of genomic instability (Supplementary Table 9).

The remaining half of the insertional bias (~1.5 Mb) was accounted for by 1,116 more complex structural variants (which we define as insertions having either several annotated repeat elements, or at least 30% of

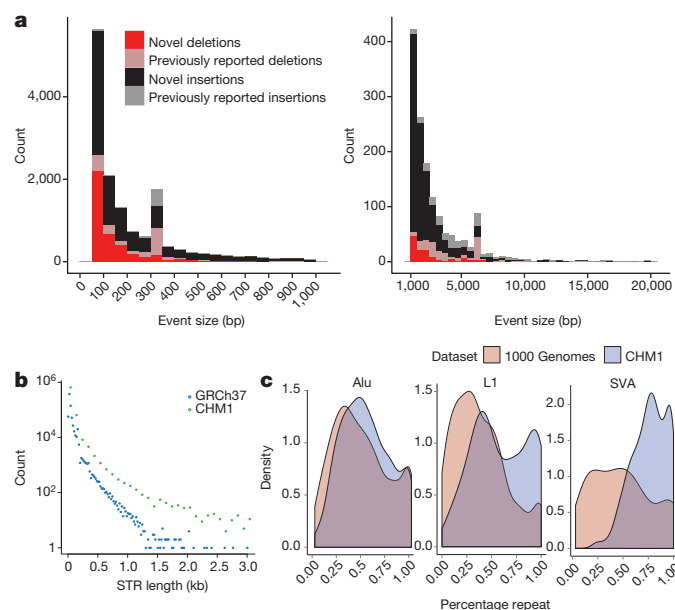


Figure 2 | Structural variation analyses. **a**, Histograms display the distribution of novel insertions (black/grey) and deletions (red/pink) between CHM1 and GRCh37 haplotypes compared to copy number variants identified from other studies for insertions and deletions less than 1 kb (left) and greater than or equal to 1 kb (right). Most of the increased sensitivity occurs below 5 kb. Peaks at ~300 bp and 6 kb correspond to Alu and L1 insertions, respectively. **b**, STR insertions in CHM1 (green) are longer than the human genome (blue; GRCh37), and this effect becomes more pronounced with increasing length (x axis). **c**, The percentage repeat composition (x axis) of 1-kb sequences flanking insertion sites for Alu, L1 and SVA mobile element insertions. Insertion calls from the 1000 Genomes Project (pink)²¹ compared to calls from CHM1 using SMRT reads (blue) show increased sensitivity for repeat-rich insertions.

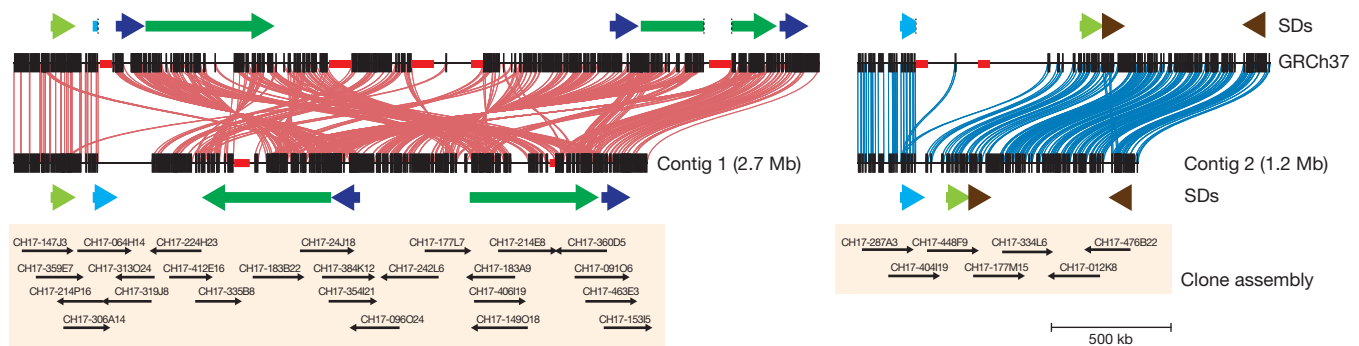


Figure 3 | CHM1 clone-based assembly of the human 10q11 genomic region. The clone-based assembly is composed primarily of BACs from the CH17 library as shown in the tiling path below the internal repeat structure of

the remaining sequence not annotated as repeat) (Table 1 and Extended Data Fig. 4). Sequence analyses of these regions of the genome revealed these insertions were frequently embedded within regions already enriched for clusters of mobile element insertions. Complex repetitive regions such as these represent a major challenge in structural variant detection owing to spurious mapping of short-read sequence data. We performed site complexity analysis of annotated mobile element insertion loci by assessing the repeat composition of the 1-kb sequences 5' and 3' flanking the retrotransposons AluY, L1 and SVA insertions in both the CHM1 sequencing data and insertion sites from population-scale low-coverage sequencing data²¹. While we observed a small bias in the repeat complexity of AluY insertions (53% versus 48%; $P = 4.8 \times 10^{-6}$, Kolmogorov–Smirnov test), a much more marked shift is seen for L1 and SVA insertions. We found that human-specific L1Hs insertion sites in CHM1 have a flanking common repeat content of 59% when compared to 39% in the 1000 Genomes Project data set ($P = 1.8 \times 10^{-10}$, Kolmogorov–Smirnov test) (Fig. 2c). The bias for SVA insertions is even greater, with 76% of insertions mapping adjacent to repeats when compared to 50% using Illumina read-pair data ($P = 3.84 \times 10^{-14}$, Kolmogorov–Smirnov test).

The large STR and complex insertions are enriched for regions annotated as having potential clone assembly problems. This enrichment becomes more pronounced the larger and more complex the insertion (for example, the 185-fold enrichment of ‘black tag’ annotations for STR insertions; Supplementary Information). Notably, less than 1% of these variants are present in newer assemblies of the human genome, including GRCh38 and CHM1.1 (ref. 22) (derived primarily by Illumina sequencing technology). Because we find evidence of most of these complex events in additional human or chimpanzee genomes (Supplementary Information), we propose that ~1,700 sites (3.5 Mb) represent deficiencies or ‘muted’ gaps that can now be accessed as a result of SMRT technology (Supplementary Table 7). We incorporated these inserted sequences as well as gap closures into a patched GRCh37 reference, effectively mapping 0.026% additional Illumina reads and discovering additional single nucleotide polymorphisms (SNPs) (for example, 9,231 SNPs; Supplementary Information).

In addition to insertions and deletions, we also searched for the presence of inversions—a structural variation class that is notoriously difficult to ascertain. We developed a search algorithm that specifically leveraged the increased length of the SMRT sequence reads to search for ‘reversals’ in order when aligned to the reference. Regions with two or more reversals were then locally assembled to define the breakpoints of each event optimally. We identified 34 inversions with an average length of 7.1 kb, corresponding to a total of ~240 kb of inverted sequence (Supplementary Table 8 and Supplementary Fig. 6). We subcloned and sequenced 15 events using a large-insert BAC library with a validation rate of 100% (15 out of 15) (Extended Data Fig. 5). None of the events disrupted genes, no enrichment was observed on the X chromosome,

the region. Coloured arrows indicate large segmental duplications (SDs) with homologous sequences connected by lines generated by Miropeats²³.

and 68% (23 out of 34) of the inversions were flanked by inverted repeats (Supplementary Table 8).

A limitation of our approach is its dependence on the local assembly of mapped reads to the human reference genome. Even with an average mapped read length of 5.8 kb, not all reads may be uniquely mapped to a specific location. As a result, gaps ($n = 82$) adjacent to segmental duplications were largely unresolved, inversions exceeding the read length (>20 kb) could not be detected (for example, 15q13.3 region), and SMRT sequence read synthesis within or flanking long, highly identical repeats could not be reliably assembled. We identified a total of 737 euchromatic regions (12.5 Mb) of our genome, in which large-scale mapping inconsistencies ($n = 22$) or deficiencies ($n = 715$) were noted but were unresolvable by this approach (Supplementary Tables 26 and 27). We selected one 6.5-Mb region mapping to chromosome 10q11.23 for a more detailed analysis. The region carried seven gaps within the human reference genome (GRCh37), none of which was resolved or extended by SMRT WGS reads. We applied an alternative clone-based hierarchical approach (Supplementary Information) and identified a tiling path of 32 BACs and assembled the clone inserts using SMRT sequencing¹⁴. We generated sequence contigs spanning two large clusters of segmental duplication (2.7 and 1.2 Mb), closing six of the seven gaps in this region (Fig. 3 and Extended Data Fig. 6), adding 416 kb of missing reference sequence, correcting the orientation of 1,451 kb, and eliminating 856 kb of redundant sequence that was represented twice within the reference. Two gaps remain, each at the same location within paralogous segmental duplications, corresponding to a nearly perfect 50-kb tandem repeat that cannot be resolved at the level of large-insert clones using existing methods. These results indicate that although it is possible to use reads to close gaps and detect variation missed by other next-generation sequencing methods, the resolution of larger, complex regions of the genome still require targeted efforts that leverage both clones and WGS data. Complete *de novo* assembly of human genomes will probably require the development of even longer-range sequencing data. The approaches outlined here will have broader application to many of the unfinished and complex regions of mammalian genomes.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 3 July; accepted 30 September 2014.

Published online 10 November 2014.

1. The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
2. The International HapMap Project Consortium. The International HapMap Project. *Nature* **426**, 789–796 (2003).
3. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931–945 (2004).
4. Kurahashi, H. *et al.* Molecular cloning of a translocation breakpoint hotspot in 22q11. *Genome Res.* **17**, 461–469 (2007).

5. Genovese, G. *et al.* Using population admixture to help complete maps of the human genome. *Nature Genet.* **45**, 406–414 (2013).
6. Bovee, D. *et al.* Closing gaps in the human genome with fosmid resources generated from multiple individuals. *Nature Genet.* **40**, 96–101 (2008).
7. Mills, R. E. *et al.* Mapping copy number variation by population-scale genome sequencing. *Nature* **470**, 59–65 (2011).
8. Kidd, J. M. *et al.* A human genome structural variation sequencing resource reveals insights into mutational mechanisms. *Cell* **143**, 837–847 (2010).
9. Eichler, E. E., Clark, R. A. & She, X. An assessment of the sequence gaps: unfinished business in a finished human genome. *Nature Rev. Genet.* **5**, 345–354 (2004).
10. Eid, J. *et al.* Real-time DNA sequencing from single polymerase molecules. *Science* **323**, 133–138 (2009).
11. Chaisson, M. J. & Tesler, G. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. *BMC Bioinformatics* **13**, 238 (2012).
12. Lee, H. & Schatz, M. C. Genomic dark matter: the reliability of short read mapping illustrated by the genome mappability score. *Bioinformatics* **28**, 2097–2105 (2012).
13. Myers, E. W. *et al.* A whole-genome assembly of *Drosophila*. *Science* **287**, 2196–2204 (2000).
14. Chin, C. S. *et al.* Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature Methods* **10**, 563–569 (2013).
15. Huddleston, J. *et al.* Reconstructing complex regions of genomes using long-read sequencing technology. *Genome Res.* **24**, 688–696 (2014).
16. Kimelman, A. *et al.* A vast collection of microbial genes that are toxic to bacteria. *Genome Res.* **22**, 802–809 (2012).
17. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
18. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
19. Conrad, D. F. *et al.* Origins and functional impact of copy number variation in the human genome. *Nature* **464**, 704–712 (2010).
20. Kong, A. *et al.* A high-resolution recombination map of the human genome. *Nature Genet.* **31**, 241–247 (2002).
21. Stewart, C. *et al.* A comprehensive map of mobile element insertion polymorphisms in humans. *PLoS Genet.* **7**, e1002236 (2011).
22. Steinberg, K. M. *et al.* Single haplotype assembly of the human genome from a hydatidiform mole. *Genome Res.* (in press).
23. Parsons, J. D. Miropeats: graphical DNA sequence comparisons. *Comput. Appl. Biosci.* **11**, 615–619 (1995).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. Alexander, D. Church and A. Klammer for discussions, K. Mohajeri and L. Harshman for technical assistance and T. Brown for assistance in manuscript preparation. This work was supported, in part, by US National Institutes of Health (NIH) grant HG002385 and HG007497 to E.E.E. M.Y.D. is supported by the US National Institute of Neurological Disorders and Stroke (award K99NS083627). E.E.E. is an investigator of the Howard Hughes Medical Institute.

Author Contributions E.E.E., M.J.P.C., M.Y.D., J.H. and J.K. designed experiments; M.M. prepared DNA; M.M. and M.B. prepared libraries and generated sequence data; P.H.S., J.H. and M.Y.D. identified clones for sequencing; J.H., P.H.S., M.Y.D., F.H. and M.J.P.C. performed bioinformatics analyses; M.Y.D., F.A. and M.M. performed targeted sequencing of clones; M.J.P.C. designed algorithms and pipelines for mapping SMRT sequence data and detection of structural variants; M.W.H., U.S., R.S. and J.A.S. provided access to critical resources; J.M.L. deposited SMRT sequence data into SRA; M.J.P.C., J.H. and E.E.E. wrote the manuscript.

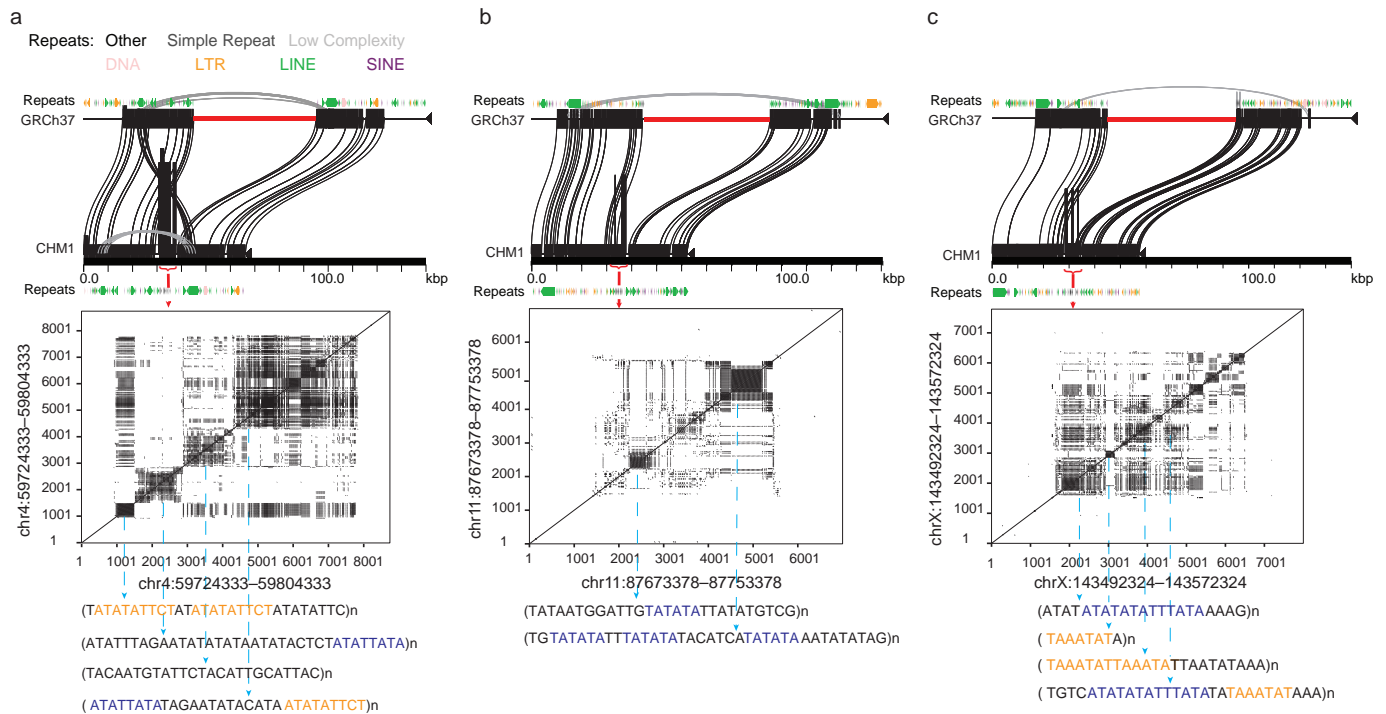
Author Information All underlying SMRT WGS read data have been released within the NCBI Sequence Read Archive (SRA) under accession SRX533609 and may also be accessed as part of all the SMRT data sets (NCBI SRA accession SRP040522). Illumina WGS data for CHM1 are available in the NCBI SRA under accession SRP044331 as well as finished BAC and fosmid clone inserts using SMRT sequence data (GenBank accessions in Supplementary Table 35). For the purpose of mapping and annotation, we developed a patched GRCh37 reference genome including a track hub for upload into the UCSC Genome Browser. A complete list of all inaccessible regions of the human genome and a database of heterochromatic and subtelomeric sequence reads that could not be assembled are available at (<http://eichlerlab.gs.washington.edu/publications/chm1-structural-variation>). Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.E.E. (eee@gs.washington.edu).

METHODS

SMRT WGS data (41-fold sequence coverage) was generated using a Pacific Biosciences RSII instrument (P5C3 chemistry) from genomic libraries generated from a complete hydatidiform mole DNA (CHM1tert). Sequence reads were mapped to the human reference genome (GRCh37) using a modified version of BLASR (<http://www.github.com/EichlerLab/blasr>) (Supplementary Methods); a bioinformatics pipeline was developed to identify regions of structural variation and extensions into gaps (http://www.github.com/EichlerLab/chm1_scripts); corresponding sequence reads were *de novo* assembled and a high-quality consensus sequence generated for each region using Celera v.8.1 (ref. 13) and Quiver v.0.7.6 (ref. 14). Reads are selected for support of a variant if the mapping quality is greater than 20; a minimum of 5 reads are required to trigger an assembly. For the purpose of this analysis, we focused only on the euchromatic portion of the genome excluding pericentromeric regions (5 Mb flanking annotated centromeres), all acrocentric portions of chromosomes, and subtelomeric regions (150 kb from the annotated telomeric sequence). Repeat content of all structural variants was determined using CENSOR²⁴, RepeatMasker²⁵, Miropeats²³ and TRF (<http://tandem.bu.edu/>). The sequence accuracy of the assemblies

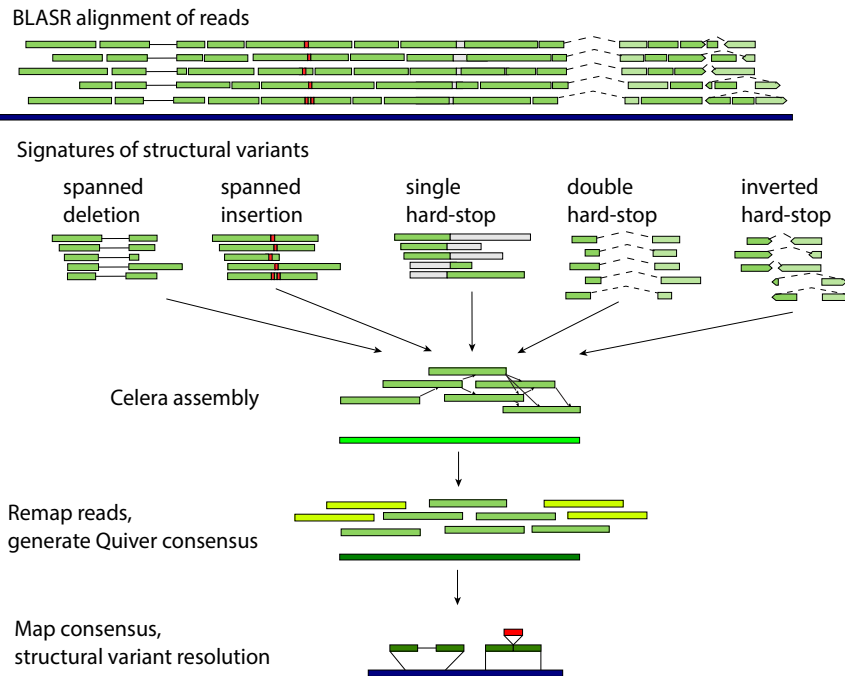
and structural variant polymorphisms were inferred by comparison to 18 sequenced large-insert BAC (CH17) and 89 fosmid clones⁸, Sanger-based BAC-end sequence generated for CHM1tert (GenBank accessions in Supplementary Table 35), and comparison to Illumina-based WGS generated for human genomes¹. We also generated Illumina WGS data (41-fold) for comparison (SRA SRP044331). For the chromosome 10q11 region, 125 CH17 BACs were identified and sequenced using a Nextera-Illumina protocol²⁶. A minimal tiling path of 35 clones was deeply sequenced (300-fold coverage) using 1 SMRT cell per clone; inserts were assembled and an alternative reference was created using methods described previously¹⁵.

24. Jurka, J., Klonowski, P., Dagman, V. & Pelton, P. CENSOR—a program for identification and elimination of repetitive elements from DNA sequences. *Comput. Chem.* **20**, 119–121 (1996).
25. Smit, A. F. A., Hubley, R. & Green, P. RepeatMasker Open-3.0 <http://www.repeatmasker.org> (1996–2010).
26. Adey, A. *et al.* Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol.* **11**, R119 (2010).
27. Wu, T. & Watanabe, C.K. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**, 1859–1875 (2005).



Extended Data Figure 1 | Sequence content of gap closures. **a–c**, Gap closures are enriched for simple repeats compared to equivalently sized regions randomly sampled from GRCh37; examples of the organization of these regions are shown using Miropeats for chromosome 4 (GRCh37, chr4:59724333–59804333) (**a**), chromosome 11 (GRCh37,

chr11:87673378–87753378) (**b**), and chromosome X (GRCh37, chrX:143492324–143572324) (**c**). Dotplots show the architecture of the degenerate STRs with the core motif highlighted below. Shared sequence motifs between blocks are indicated by colour.



Extended Data Figure 2 | Variant detection pipeline. At every variant locus, we collected the full-length reads that overlap the locus, performed *de novo* assembly using the Celera assembler, and called a consensus using Quiver after remapping reads used in the assembly as well as reads flanking the assembly (yellow reads) to increase consensus quality at the boundaries of the assembly. BLASR is used to align the assembly consensus sequences to the reference,

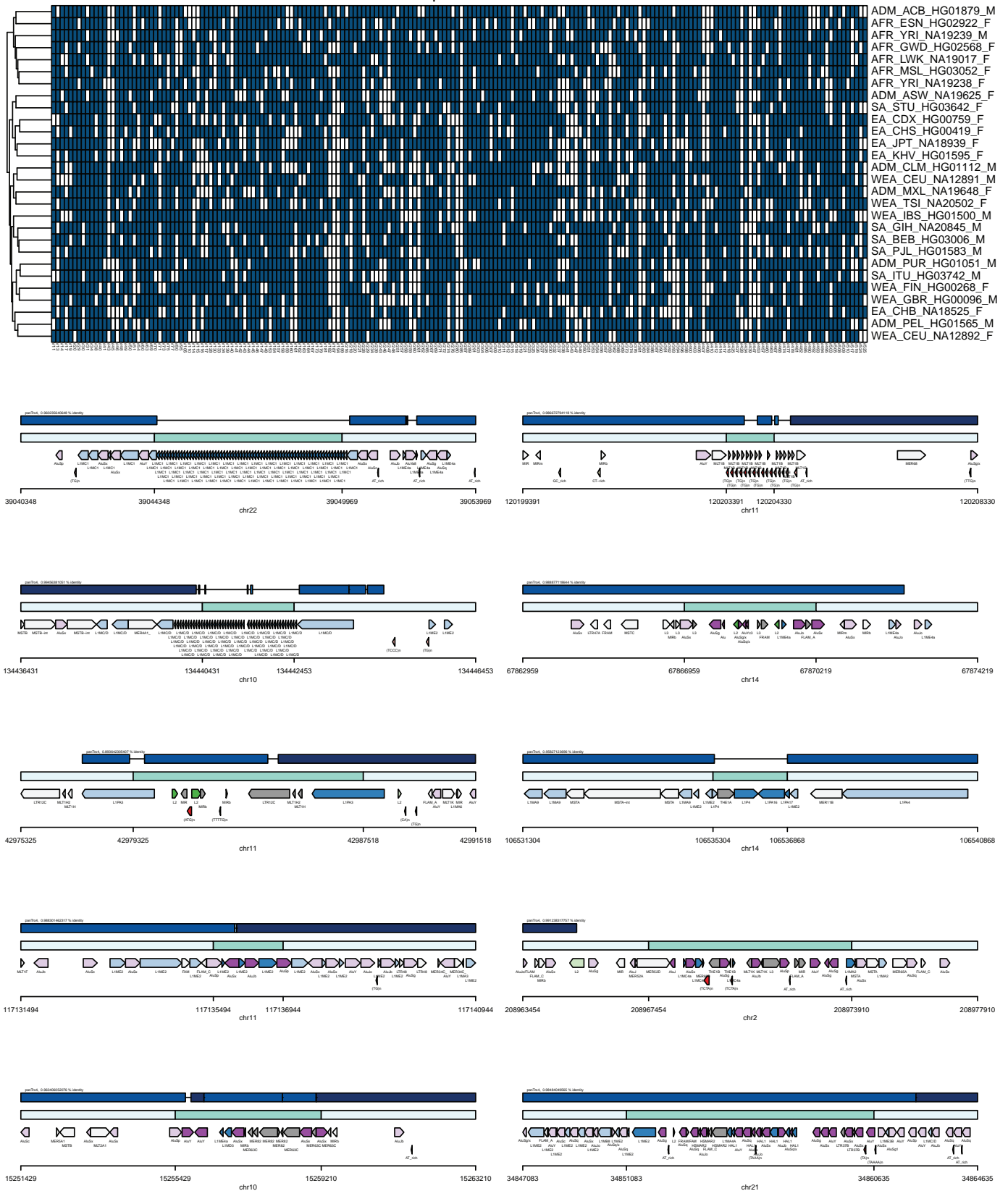
and insertions and deletions in the alignments are output as variants. Reads spanning a deletion event within a single alignment are shown as bars connected by a solid line, and double hard-stop reads spanning a larger deletion event and split into two separate alignments of the same read are shown as a dotted line.



Extended Data Figure 3 | Genome distribution of closed gaps and insertions. Chromosome ideogram heatmap depicts the normalized density of inserted CHM1 base pairs per 5-Mb bin with a strong bias noted near the end of

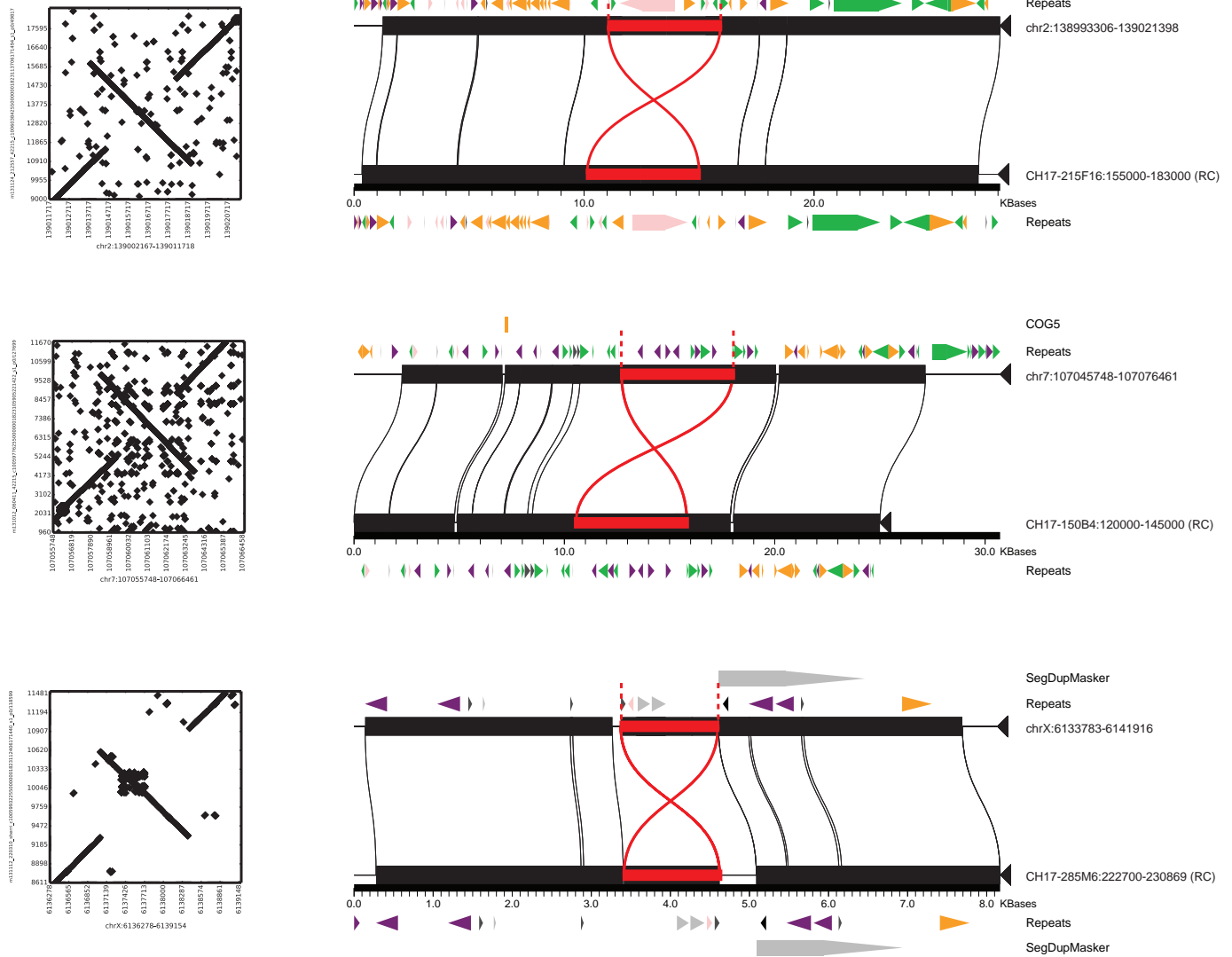
most chromosomes. Locations of structural variants and closed gaps are given by coloured diamonds to the left of each chromosome: closed gap sequences (red), inversions (green), and complex events (blue).

Complex insertions



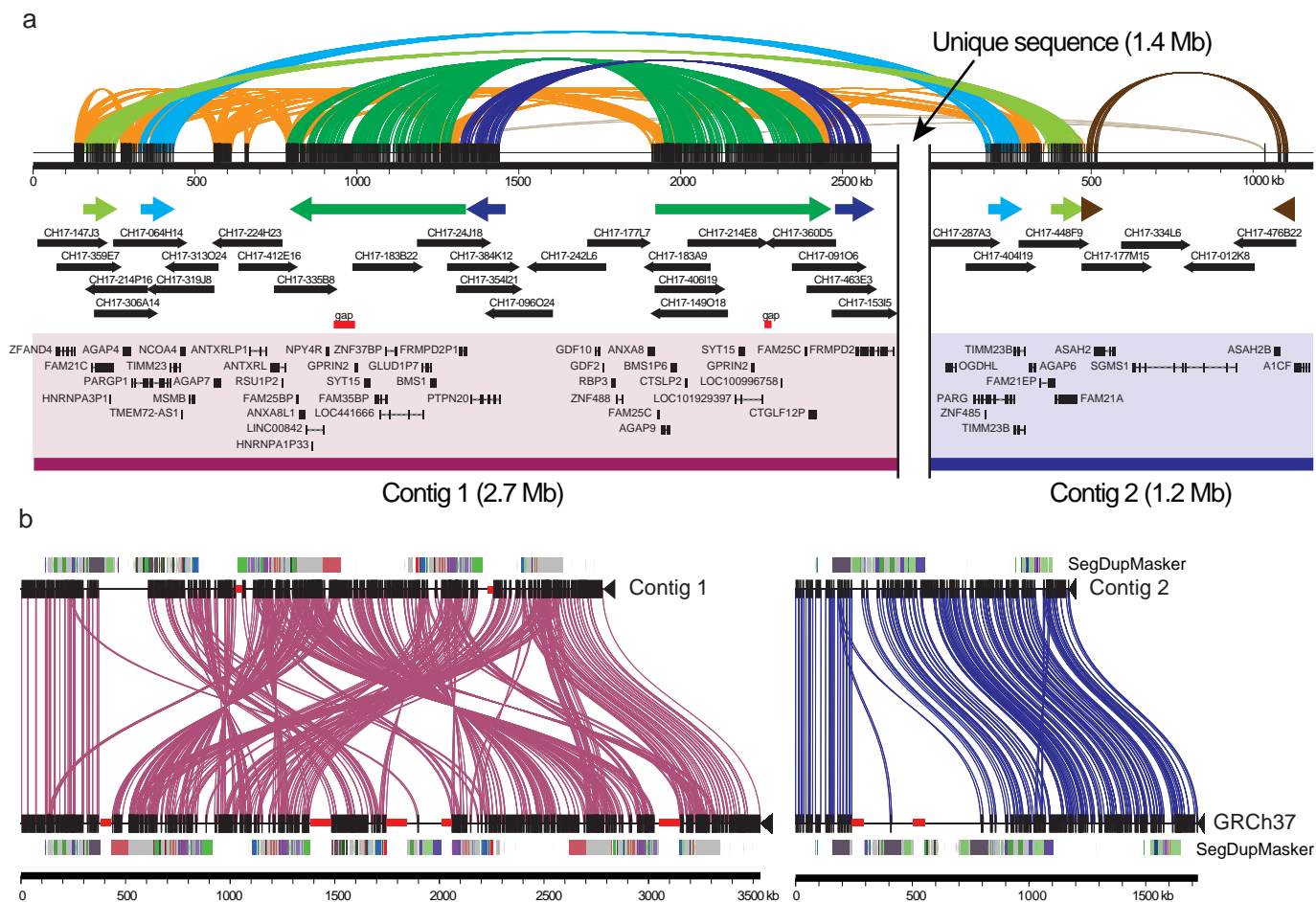
Extended Data Figure 4 | Confirmation of complex insertions in additional genomes. Top, genotypes of polymorphic complex regions using read depth of unique k -mers (blue: present; white: absent). Bottom, extended examples of complex insertion events: alignment to chimpanzee panTro4 reference

(dark blue); existing human reference hg19 (light teal); inserted sequence (dark teal). The bottom rows show repeat annotations, with darker hues for repeats overlapping the inserted region.



Extended Data Figure 5 | Inversion validation by BAC-insert sequencing. Inversions detected by alignment of single long reads were validated by sequencing clones from the CHM1 BAC library (CHORI17), in which end mappings to GRCh37 spanned the putative inversions. Inversions were

validated by aligning the corresponding BAC sequences to GRCh37 with Miropeats. Shared sequence between the BACs and GRCh37 is shown in black; inversion events are indicated in red.



Extended Data Figure 6 | CHM1 clone-based assembly of the human 10q11 genomic region. **a**, The clone-based assembly is composed primarily of BACs from the CH17 library as shown in the tiling path below the internal repeat structure of the region. Coloured arrows indicate large segmental duplications with homologous sequences connected by coloured lines (Miropeats). Genes

annotated from alignment of RefSeq messenger RNA sequences with GMAP²⁷ are shown. **b**, Miropeats comparisons of the 10q11 clone-based assembly against the corresponding sequence from GRCh37, with gaps shown in red, highlight the degree to which the reference was misassembled.

Interception of host angiogenic signalling limits mycobacterial growth

Stefan H. Oehlers¹, Mark R. Cronan¹, Ninecia R. Scott¹, Monica I. Thomas¹, Kazuhide S. Okuda², Eric M. Walton¹, Rebecca W. Beerman¹, Philip S. Crosier² & David M. Tobin¹

Pathogenic mycobacteria induce the formation of complex cellular aggregates called granulomas that are the hallmark of tuberculosis^{1,2}. Here we examine the development and consequences of vascularization of the tuberculous granuloma in the zebrafish–*Mycobacterium marinum* infection model, which is characterized by organized granulomas with necrotic cores that bear striking resemblance to those of human tuberculosis². Using intravital microscopy in the transparent larval zebrafish, we show that granuloma formation is intimately associated with angiogenesis. The initiation of angiogenesis in turn coincides with the generation of local hypoxia and transcriptional induction of the canonical pro-angiogenic molecule Vegfaa. Pharmacological inhibition of the Vegf pathway suppresses granuloma-associated angiogenesis, reduces infection burden and limits dissemination. Moreover, anti-angiogenic therapies synergize with the first-line anti-tubercular antibiotic rifampicin, as well as with the antibiotic metronidazole, which targets hypoxic bacterial populations³. Our data indicate that mycobacteria induce granuloma-associated angiogenesis, which promotes mycobacterial growth and increases spread of infection to new tissue sites. We propose the use of anti-angiogenic agents, now being used in cancer regimens, as a host-targeting tuberculosis therapy, particularly in extensively drug-resistant disease for which current antibiotic regimens are largely ineffective.

The human tuberculous granuloma, a tightly cohesive cellular structure that houses infecting mycobacteria, develops hypoxic areas around its necrotic core⁴. In tumours, the development of hypoxia is tightly linked to angiogenesis and subsequent metastasis⁵. In tuberculosis, attention has focused on the possible consequences of granuloma hypoxia to bacterial physiology⁴, but relatively little attention has been paid to the functional significance of findings that tuberculous granulomas are extensively vascularized^{6–8}.

In its natural ectothermic hosts, *M. marinum*, the closest relative of the *M. tuberculosis* complex, causes a disease called fish tuberculosis, a

systemic wasting disease with organized epithelioid granulomas with necrotic cores. In zebrafish larvae, mycobacterium-infected macrophages form early granulomas, undergo a hallmark epithelioid transformation, and activate granuloma-specific gene expression programs^{2,9}.

To monitor host vasculature in zebrafish, we used the *Tg(kdr:egfp)*⁸⁴³ line (referred to hereafter as *Tg(flk1:eGFP)*), in which vascular endothelial cells are fluorescently labelled with enhanced green fluorescent protein (eGFP)¹⁰. Injection of mycobacteria into the most commonly used caudal vein site results in granulomas in the immediate vicinity of the richly vascularized area of the caudal haematopoietic tissue (CHT) (Fig. 1a). To determine whether a different injection site with sparser and smaller blood vessels was more suitable to detect angiogenesis, we assessed primary granulomas that typically formed dorsally after injection into the trunk (Fig. 1b, c). As with caudal vein injection, trunk injection resulted in most bacteria becoming resident in macrophages in granulomas (Extended Data Fig. 1a). Trunk granulomas progressed similarly to caudal vein granulomas, with dissemination into richly vascularized areas¹¹ (Supplementary Video 1). Using *Tg(mpeg1:tdTomato-caax^{xt3})* larvae, in which macrophages are labelled by membrane-bound Tomato, revealed similar macrophage dynamics¹¹, including the interstitial egress of infected macrophages, the transfer of *M. marinum* between granulomas, and coalescence of distal bacteria into existing granulomas (Supplementary Videos 2, 3, 4). Some infected macrophages invaded vasculature around the primary granuloma (Extended Data Fig. 1b and Supplementary Video 5).

Imaging of trunk-infected larvae revealed angiogenesis, with growth of vasculature around sites of infection. We observed sprouting from the existing intersegmental vessels (ISVs) starting at 4 days post-infection (dpi), just after the formation of granulomas (Fig. 1c). We assessed the features of vessel sprouting through long-term live imaging of infected larvae (Supplementary Videos 6, 7, 8, 9). Vessel growth occurred in spurts

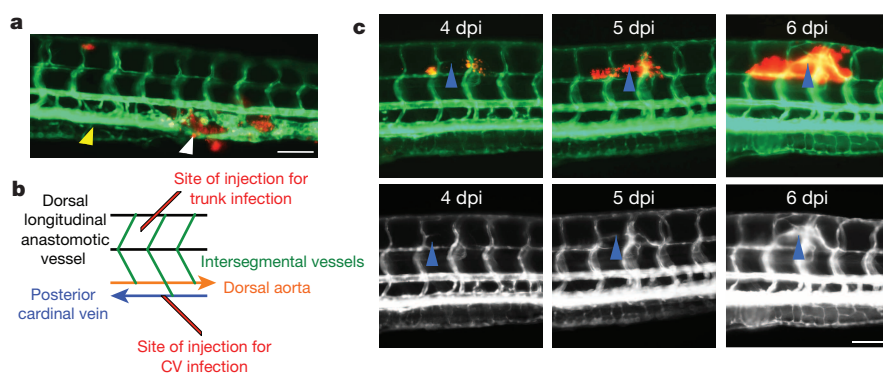


Figure 1 | *M. marinum* infection induces angiogenesis in the zebrafish infection model. **a**, Tomato-fluorescent *M. marinum* granuloma in the CHT region of a *Tg(flk1:eGFP)* larva. White arrowhead indicates area of occlusion in the posterior cardinal vein caused by *M. marinum* granuloma. Yellow arrowhead indicates area of normal posterior cardinal vein width anterior of occlusion. **b**, Schematic depicting location of injection into the trunk of a 2 days post-fertilization larva. CV, caudal vein. **c**, Time-lapse images of vascular growth around a trunk granuloma from a single *Tg(flk1:eGFP)* larva from 4 dpi to 6 dpi. Top, Tomato-fluorescent *M. marinum* and labelled vasculature. Bottom, only *Tg(flk1:eGFP)*-labelled vasculature. Blue arrowhead tracks the growth of a single vessel across all frames. Images are representative of 10 (**a**) and 20 (**c**) individual animals. Scale bars, 100 μ m.

¹Department of Molecular Genetics and Microbiology, Center for Microbial Pathogenesis, Duke University Medical Center, Durham, North Carolina 27710, USA. ²Department of Molecular Medicine and Pathology, The University of Auckland, Auckland 1023, New Zealand.

with extended periods of quiescence or even reversed directionality (Extended Data Fig. 1c).

To examine the mode of vascular elongation, we used blue-fluorescent *M. marinum* to infect the transgenic zebrafish line *Tg(fli1a:nlsGFP)⁹⁷; flk1:mCherry¹⁵⁵*, in which endothelial nuclei are marked by eGFP expression. At 4 dpi, nuclei left the highly organized ISVs, always towards sites of infection, and subsequently divided within the somites (Extended Data Fig. 1d and Supplementary Video 10). Vessels sprouted from both arterial and venous ISVs (Extended Data Fig. 1e).

We determined whether new blood vessels generated around the granuloma were functional. Using DsRed-labelled erythrocytes in the transgenic line *Tg(flkl:eGFP; gata1:DsRed^{sd2})* we found substantial blood flow through both ectopic vessels that spanned existing vessels completely and into newer blind-ending vessels (Extended Data Fig. 1f).

Angiogenesis required persistent *M. marinum* infection; it did not develop after injection of PBS, heat-killed *M. marinum* or non-pathogenic *Escherichia coli* (Extended Data Fig. 2a). Tumour-associated macrophages are important drivers of tumour angiogenesis upon tumour hypoxia¹². Since macrophages serve as a principal repository of virulent mycobacteria, we assessed whether there were differences in vascular recruitment between macrophage-resident and extracellular mycobacteria. Infection of double-transgenic *Tg(flkl:eGFP; mpeg1:tdTomato-caax^{xt3})* embryos with Cerulean-fluorescent *M. marinum* allowed us to discriminate between intracellular and extracellular bacteria (Fig. 2a). Enumeration of vascular branching revealed an elevated vascularization rate for intracellular compared with extracellular foci (odds ratio (OR) and 95% confidence interval: 4 dpi, 6.63 (2.57–17.11); 5 dpi, 6.93 (1.51–31.75); 6 dpi, 27.84 (10.18–76.18)) (Extended Data Fig. 2b).

To demonstrate a functional requirement for macrophages in recruiting vasculature we performed morpholino knockdown of the transcription factor Pu.1 (also known as Spi-1), which fully ablates all macrophages until 5 days post-fertilization (dpf) (Extended Data Fig. 2c)¹³. As previously reported¹⁴, infection burden was markedly increased in the Pu.1 knockdown animals (Extended Data Fig. 2d). Despite increased burden, the total length of abnormal vasculature was decreased in morpholino-injected animals compared with controls, suggesting that macrophages specifically mediate new vessel growth in the context of inflammation (Fig. 2b).

Vascularization coincided with the formation of granulomas at around 4 dpi. We disrupted mycobacterium-driven granuloma formation using a *M. marinum* strain deficient in the ESX1 protein export system (Δ ESX1), which could not form granulomas¹⁵. We infected Δ ESX1 with a 7.5-fold excess over wild type to generate equivalent bacterial burdens at 4 dpi (Extended Data Fig. 2e). As expected, both strains were predominantly macrophage resident, but Δ ESX1 infection resulted in fewer granulomas and a marked reduction in angiogenesis (Fig. 2c). Thus, macrophage residence *per se* is not sufficient to induce angiogenesis. Rather, macrophages that have undergone further differentiation to form granulomas appear to be required.

When we analysed the relationship between the size of individual infection foci and the length of recruited vasculature we found a strong correlation (Extended Data Fig. 3a). This relationship is reminiscent of the angiogenic switch, in which tumour size is related directly to the requirement for vascularization due to the development of local hypoxia⁵.

We therefore asked whether hypoxia develops within granulomas. Activation of Hif-1 α increases transcription of *prolyl hydroxylase 3* (*phd3*; also known as *egln3*), which serves as a reporter for hypoxic conditions in zebrafish larvae^{16,17}. We detected robust expression of *phd3* within trunk granulomas, but not CHT granulomas, which are already proximate to the vasculature (Extended Data Fig. 3b), consistent with a lack of hypoxia in the CHT¹⁸. Conversely, to assess whether mycobacteria within granulomas were experiencing hypoxia, we investigated the effect of metronidazole treatment¹⁹. Metronidazole specifically kills anaerobically growing bacteria, including anaerobically growing *M. tuberculosis*, but is ineffective during aerobic growth^{3,19,20}. We used a nitroreductase-expressing transgenic line *Tg(lyzC:ntr-p2A-lanYFP^{xt14})* to titrate a

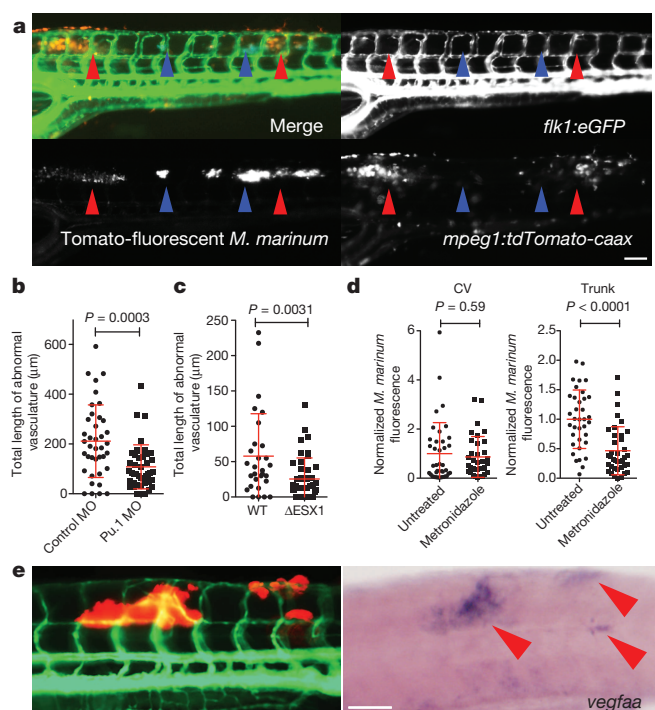


Figure 2 | *M. marinum* induces vascularization through granuloma formation in cooperation with host leukocytes and the expression of Vegf. **a**, Cerulean-fluorescent *M. marinum* distribution in a *Tg(flkl:eGFP; mpeg1:tdTomato-caax^{xt3})* double-transgenic larva. Blue arrows indicate sites of extracellular bacterial growth, red arrows indicate sites of intracellular containment. Image is representative of 48 individual animals. **b, c**, Length of abnormal vasculature in 5 dpi control and Pu.1 morphant larvae (**b**), and 4 dpi larvae infected with wild-type (WT) or Δ ESX1 Tomato-fluorescent *M. marinum* (**c**). MO, morpholino. Student's *t*-test with Welch's correction, all data are pooled from two biological replicates. **d**, Bacterial burden in caudal-vein (CV)-infected (left) and trunk-infected larvae (right) treated with 5 mM metronidazole. Student's *t*-test, data are pooled from three biological replicates. **e**, Distribution of Tomato-fluorescent *M. marinum* in a *Tg(flkl:eGFP)* larva (left) and corresponding whole-mount *in situ* hybridization detection of *vegfaa* expression in the same larva (right). Red arrowheads indicate sites of *M. marinum* granulomas. Image is representative of 20 individual animals. Scale bars, 100 μ m. Error bars represent mean \pm standard deviation (s.d.).

biologically active dose of metronidazole (Extended Data Fig. 3c). Consistent with the *phd3* staining, metronidazole treatment reduced bacterial burden in trunk-infected animals but not caudal-vein-infected larvae (Fig. 2d).

In tumours, hypoxia is known to induce VEGF expression, which in turn stimulates angiogenesis²¹. VEGF has been associated with tuberculosis pathogenesis: it is induced in active pulmonary tuberculosis²², and, in a rat corneal model, mediates neovascularization in granulomas triggered by mycobacterial trehalose-6,6'-dimycolate²³. Vegf has a conserved role in homeostatic blood vessel recruitment in zebrafish²⁴. We observed strong *vegfaa* expression around sites of mycobacterial granulomas in the trunk (Fig. 2e). Similar expression levels were also observed around CHT granulomas, indicating that initiation of pro-angiogenic signalling is not dependent on prior development of hypoxia during infection (Extended Data Fig. 4a). Analysis of stained sections revealed cells around the edge of mycobacterial granulomas expressing *vegfaa*, an observation consistent with a primarily macrophage-driven expression pattern that could be reduced by macrophage depletion (Extended Data Fig. 4b). In addition to its role in angiogenesis, VEGF plays an important part in driving vascular permeability²⁵. We performed micro-angiography on infected animals and found increased vascular leakage, suggesting a local effect of *vegfaa* expression around infection foci (Extended Data Fig. 4c).

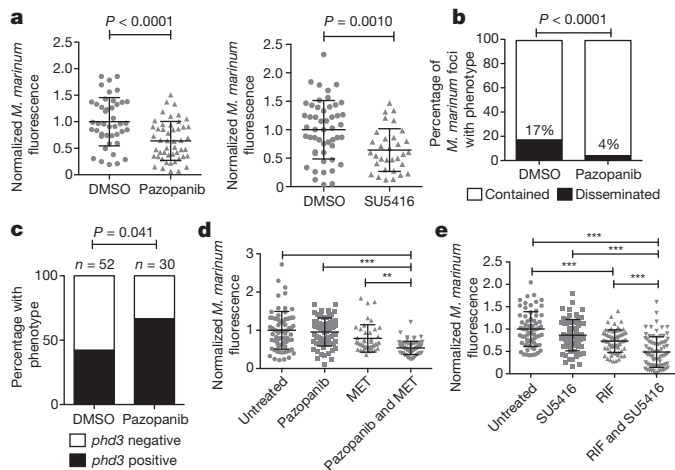


Figure 3 | Inhibition of VEGFR signalling reduces *M. marinum* pathogenicity in zebrafish larvae. **a**, Bacterial burden in trunk-infected pazopanib-treated (left) and SU5416-treated (right) larvae. Student's *t*-test, data are pooled from two (left graph) or three (right graph) biological replicates. DMSO, dimethylsulphoxide. **b**, Bacterial dissemination in untreated and pazopanib-treated larvae. Total number of granulomas and larvae analysed: untreated, 77 granulomas from 18 larvae; pazopanib, 130 granulomas from 22 larvae. Fisher's exact test. **c**, Expression of *phd3* hypoxia marker in untreated and pazopanib-treated infected larvae detected by *in situ* hybridization. Total number of larvae analysed: 52 (DMSO); 30 (pazopanib). Fisher's exact test, data are from a single technical replicate of two pooled biological replicates. **d**, Bacterial burden in pazopanib-treated, metronidazole (MET)-treated, and pazopanib and metronidazole treated larvae. One-way analysis of variance (ANOVA) with Tukey's post-test, data are pooled from three biological replicates. **e**, Bacterial burden in rifampicin (RIF)-treated, SU5416 and rifampicin treated, and SU5416-treated larvae. One-way ANOVA with Tukey's post-test, data are pooled from three biological replicates. Error bars represent mean \pm s.d. ** $P < 0.01$, *** $P < 0.001$.

Angiogenesis and, specifically, VEGFR signalling have been targeted in cancer therapies. Our findings suggested that these therapies might also be useful for mycobacterial infections. We chose the well-characterized small molecule SU5416, a prototypical tyrosine kinase receptor, and pazopanib, a clinically relevant VEGFR inhibitor^{26,27}. Treatment of infected animals with SU5416 or pazopanib prevented ectopic angiogenesis around the forming granulomas in *Tg(flk1:eGFP)* zebrafish and reduced net bacterial burdens (Fig. 3a and Extended Data Fig. 5a). Neither compound affected *in vitro* growth of *M. marinum*, suggesting

that the effect on bacterial burden was achieved through targeting host pathways (Extended Data Fig. 5b). We confirmed that the treatments were specifically targeting angiogenesis, since bacterial burdens were lowered only in trunk-infected and not in caudal-vein-infected zebrafish (Extended Data Fig. 5c). Additionally, in trunk-infected animals, growth restriction did not occur until after the initiation of angiogenesis at 4 dpi (Extended Data Fig. 5d). To determine whether VEGFR inhibition affected macrophage recruitment, we compared the association of bacterial foci with macrophages between control and pazopanib-treated larvae. We did not observe any differences in the proportion of macrophage-associated foci at time points between 4 and 6 dpi (Extended Data Fig. 5e). Together, these data suggest that VEGFR inhibition reduces bacterial burden specifically through restriction of vascularization.

To determine whether VEGFR inhibitors also reduced infection-induced vascular permeability, we measured permeability in treated and untreated animals matched for infection burden. We found a reduction in vascular leakiness in zebrafish treated with a VEGFR inhibitor (Extended Data Fig. 5f).

Angiogenesis is thought to have an important role in tumour metastasis⁵. Analogously, in long-term monitoring experiments, we observed a decreased rate of *M. marinum* dissemination to distal sites in zebrafish treated with a VEGFR antagonist compared with untreated animals (OR 0.27, 95% confidence interval 0.12–0.63) (Fig. 3b and Extended Data Fig. 5g). To investigate whether decreased dissemination was solely a consequence of reduced bacterial growth, we examined a hypoinflammatory state caused by knockdown of an enzyme involved in eicosanoid biosynthesis, *Lta4h*²⁸. *lta4h* knockdown increased burden, limited angiogenesis and also decreased dissemination, suggesting a role for angiogenesis in dissemination independent of burden (Extended Data Fig. 5h).

We hypothesized that reduced bacterial growth due to decreased oxygen availability may contribute to overall reduced burdens. Pazopanib treatment resulted in an increased number of *phd3*-positive granulomas (Fig. 3c and Extended Data Fig. 5i). Moreover, pazopanib treatment increased the effectiveness of metronidazole (Fig. 3d). These results suggest that angiogenesis is an important modulator of oxygen availability for infecting mycobacteria and that its limitation can enhance the efficacy of therapies targeting hypoxia. Metronidazole has only marginal therapeutic efficacy in human tuberculosis²⁹; our results suggest combining it or related compounds with VEGFR inhibitors that increase the hypoxic environment. Finally, we showed that targeting of VEGFR signalling could complement the first-line antitubercular drug rifampicin: a combination of rifampicin and SU5416 resulted in decreased burden compared to either drug alone (Fig. 3e).

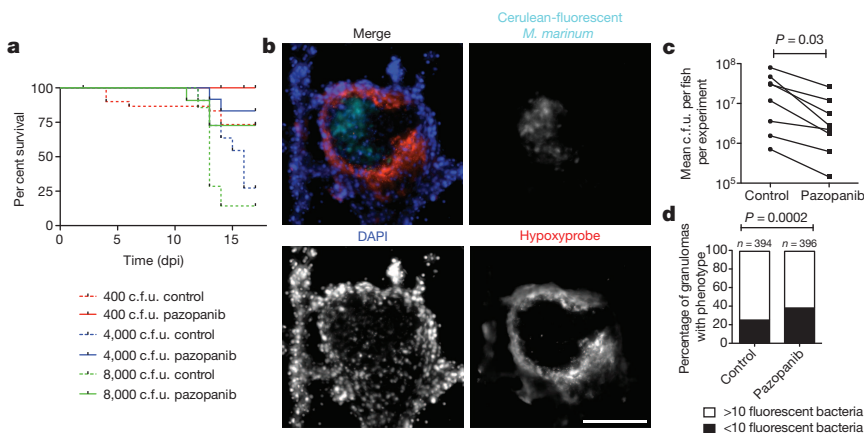


Figure 4 | Inhibition of VEGFR signalling reduces *M. marinum* burden in adult zebrafish. **a**, Survival analysis of adult zebrafish infected with 400 c.f.u. (red lines), 4,000 c.f.u. (blue lines) or 8,000 c.f.u. (green lines) of *M. marinum*. Zebrafish are further grouped into control (dashed lines) or pazopanib-treated (solid lines) groups. Log-rank test: 400 c.f.u., not significant; 4,000 c.f.u., $P = 0.012$; 8,000 c.f.u., $P = 0.029$. **b**, Representative image of a necrotic granuloma from a 2 weeks post-infection adult *Tg(flk1:eGFP)* zebrafish infected with Cerulean-fluorescent *M. marinum* (cyan), and stained for hypoxypore (red) and with 4',6-diamidino-2-phenylindole (DAPI; blue). Image is representative of granulomas found in 16 individual animals. **c**, Pooled bacterial burden in pazopanib-treated adult zebrafish. Matched Student's *t*-test. **d**, Comparison of granulomas between control and pazopanib-treated adult zebrafish scored for *M. marinum* burden as less than ten or more than ten bacteria. Total number of zebrafish analysed: 4 (control), 4 (pazopanib). Scale bar, 100 μ m. Error bars represent mean \pm s.d.

We next addressed the therapeutic effectiveness of VEGFR inhibition in adult animals. We infected zebrafish with a range of mycobacterial doses from 400–8,000 colony-forming units (c.f.u.) via intraperitoneal injection, treated them with pazopanib, and observed their survival. Over 17 days, pazopanib treatment increased survival in animals infected with high doses (4,000 and 8,000 c.f.u.) of *M. marinum*, but there was not yet appreciable mortality in the low-dose infection group (Fig. 4a). At a dose of 500 c.f.u., where significant mortality was not observed in the first 3 weeks, we observed granulomas that were completely cellular as well as ones that had developed necrotic cores by 2 weeks post-infection (wpi) (Extended Data Fig. 6a). Many adult granulomas were hypoxic, as assessed by pimonidazole treatment and staining. Staining was largely specific to granulomas and was concentrated in the cellular rim of the necrotic area, a pattern similar to that seen in macaques¹⁹ (Fig. 4b and Extended Data Fig. 6b).

Pazopanib treatment increased the mean distance of granulomas to the nearest vasculature (Extended Data Fig. 6c). As in larvae, treatment resulted in reduced bacterial burdens, with a mean fourfold reduction relative to control animals over 2 weeks (Fig. 4c). In week-old established infections, pazopanib treatment for 1 week resulted in a mean eightfold reduction in burden (Extended Data Fig. 6d).

After 6 weeks of treatment, vascularization was still significantly reduced relative to controls (Extended Data Fig. 7a). Notably, there was an increased fraction of low-burden or sterile granulomas in the pazopanib-treated zebrafish (Fig. 4d and Extended Data Fig. 7b). In addition, the drug-treated animals displayed an increased fraction of hypoxic granulomas, and there was an association of hypoxic granulomas with low-burden lesions (Extended Data Fig. 7c, d). Many low-burden or sterilized hypoxic granulomas in the drug-treated zebrafish had acellular necrotic central areas (Extended Data Figs 6a and 7b). Studies of caseous tuberculous granulomas that are sterile in asymptomatic humans suggest that such an outcome is possible even in the normal course of tuberculosis³⁰.

We conclude that angiogenesis triggered by mycobacterial granulomas is an important feature of mycobacterial pathogenesis and has important consequences for infection pathology and progression. We have shown that interception of this angiogenic program using host-directed therapies can limit mycobacterial disease. These findings suggest the potential utility of host-targeting anti-angiogenic agents as adjunctive therapies.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 November 2013; accepted 14 October 2014.

Published online 24 November 2014.

- Ernst, J. D. The immunological life cycle of tuberculosis. *Nature Rev. Immunol.* **12**, 581–591 (2012).
- Ramakrishnan, L. Revisiting the role of the granuloma in tuberculosis. *Nature Rev. Immunol.* **12**, 352–366 (2012).
- Freeman, C. D., Klutman, N. E. & Lamp, K. C. Metronidazole. A therapeutic review and update. *Drugs* **54**, 679–708 (1997).
- Rittershaus, E. S., Baek, S. H. & Sasseti, C. M. The normalcy of dormancy: common themes in microbial quiescence. *Cell Host Microbe* **13**, 643–651 (2013).
- Folkman, J. Role of angiogenesis in tumor growth and metastasis. *Semin. Oncol.* **29**, 15–18 (2002).
- Tsai, M. C. et al. Characterization of the tuberculous granuloma in murine and human lungs: cellular composition and relative tissue oxygen tension. *Cell. Microbiol.* **8**, 218–232 (2006).
- Aly, S. et al. Interferon- γ -dependent mechanisms of mycobacteria-induced pulmonary immunopathology: the role of angiostasis and CXCR3-targeted chemokines for granuloma necrosis. *J. Pathol.* **212**, 295–305 (2007).
- Ulrichs, T. et al. Differential organization of the local immune response in patients with active cavitary tuberculosis or with nonprogressive tuberculoma. *J. Infect. Dis.* **192**, 89–97 (2005).

- Davis, J. M. et al. Real-time visualization of mycobacterium-macrophage interactions leading to initiation of granuloma formation in zebrafish embryos. *Immunity* **17**, 693–702 (2002).
- Jin, S. W., Beis, D., Mitchell, T., Chen, J. N. & Stainier, D. Y. Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. *Development* **132**, 5199–5209 (2005).
- Davis, J. M. & Ramakrishnan, L. The role of the granuloma in expansion and dissemination of early tuberculous infection. *Cell* **136**, 37–49 (2009).
- Dirkx, A. E., Oude Egbrink, M. G., Wagstaff, J. & Griffioen, A. W. Monocyte/macrophage infiltration in tumors: modulators of angiogenesis. *J. Leukoc. Biol.* **80**, 1183–1196 (2006).
- Rhodes, J. et al. Interplay of pu.1 and gata1 determines myelo-erythroid progenitor cell fate in zebrafish. *Dev. Cell* **8**, 97–108 (2005).
- Clay, H. et al. Dichotomous role of the macrophage in early *Mycobacterium marinum* infection of the zebrafish. *Cell Host Microbe* **2**, 29–39 (2007).
- Volkman, H. E. et al. Tuberculous granuloma formation is enhanced by a mycobacterium virulence determinant. *PLoS Biol.* **2**, e367 (2004).
- Aprelikova, O. et al. Regulation of HIF prolyl hydroxylases by hypoxia-inducible factors. *J. Cell. Biochem.* **92**, 491–501 (2004).
- Santhakumar, K. et al. A zebrafish model to study and therapeutically manipulate hypoxia signaling in tumorigenesis. *Cancer Res.* **72**, 4017–4027 (2012).
- Elks, P. M. et al. Hypoxia inducible factor signaling modulates susceptibility to mycobacterial infection via a nitric oxide dependent mechanism. *PLoS Pathog.* **9**, e1003789 (2013).
- Via, L. E. et al. Tuberculous granulomas are hypoxic in guinea pigs, rabbits, and nonhuman primates. *Infect. Immun.* **76**, 2333–2340 (2008).
- Lin, P. L. et al. Metronidazole prevents reactivation of latent *Mycobacterium tuberculosis* infection in macaques. *Proc. Natl Acad. Sci. USA* **109**, 14188–14193 (2012).
- Forsythe, J. A. et al. Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1. *Mol. Cell. Biol.* **16**, 4604–4613 (1996).
- Matsuyama, W. et al. Increased serum level of vascular endothelial growth factor in pulmonary tuberculosis. *Am. J. Respir. Crit. Care Med.* **162**, 1120–1122 (2000).
- Saita, N., Fujiwara, N., Yano, I., Soejima, K. & Kobayashi, K. Trehalose 6,6'-dimycolate (cord factor) of *Mycobacterium tuberculosis* induces corneal angiogenesis in rats. *Infect. Immun.* **68**, 5991–5997 (2000).
- Nasevicius, A., Larson, J. & Ekker, S. C. Distinct requirements for zebrafish angiogenesis revealed by a VEGF-A morphant. *Yeast* **1**, 294–301 (2000).
- Dvorak, H. F., Brown, L. F., Detmar, M. & Dvorak, A. M. Vascular permeability factor/vascular endothelial growth factor, microvascular hyperpermeability, and angiogenesis. *Am. J. Pathol.* **146**, 1029–1039 (1995).
- Podar, K. et al. The small-molecule VEGF receptor inhibitor pazopanib (GW786034B) targets both tumor and endothelial cells in multiple myeloma. *Proc. Natl Acad. Sci. USA* **103**, 19478–19483 (2006).
- Fong, T. A. et al. SU5416 is a potent and selective inhibitor of the vascular endothelial growth factor receptor (Flk-1/KDR) that inhibits tyrosine kinase catalysis, tumor vascularization, and growth of multiple tumor types. *Cancer Res.* **59**, 99–106 (1999).
- Tobin, D. M. et al. The *Ita4h* locus modulates susceptibility to mycobacterial infection in zebrafish and humans. *Cell* **140**, 717–730 (2010).
- Carroll, M. W. et al. Efficacy and safety of metronidazole for pulmonary multidrug-resistant tuberculosis. *Antimicrob. Agents Chemother.* **57**, 3903–3909 (2013).
- Opie, E. L. & Aronson, J. D. Tubercle bacilli in latent tuberculous lesions and in lung tissue without tuberculous lesions. *Arch. Pathol. Lab. Med.* **4**, 1–21 (1927).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. Sisk and J. Saelens for technical assistance, L. Ramakrishnan, P. Edelstein and C. Kontos for helpful discussions, L. Ramakrishnan, W. Britton and J. Coers for critical review of the manuscript, and J. Fuller, C. Gallione, E. Linney, H. Mao, S. Lee, D. Marchuk, H. Matsunami, A. Nixon, J. Perfect, J. Rawls, D. Silver, K. Smith, K. Takaki, J. Tenor and B. Uy for reagents and equipment. This work was funded by an Australian National Health and Medical Research Council CJ Martin Early Career Fellowship (S.H.O.); an American Cancer Society Postdoctoral Fellowship PF-13-223-01-MPC (M.R.C.); the Duke Summer Research Opportunities Program (N.R.S.); a Malaysian Ministry of Science and Technology and Innovation scholarship (K.S.O.); a New Zealand Ministry of Science and Innovation grant UOAX0813 (P.S.C.); the Duke University Center for AIDS Research (CFAR); a National Institutes of Health (NIH)-funded program (5P30 AI064518), and by a Mallinckrodt Scholar Award, a Searle Scholar Award, a Vallee Foundation Young Investigator Award and an NIH Director's New Innovator Award 1DP2-OD008614 (D.M.T.).

Author Contributions S.H.O. and D.M.T. designed the experiments and wrote the paper. S.H.O., N.R.S., M.I.T. and K.S.O. performed and analysed the experiments. M.R.C., E.M.W. and R.W.B. generated transgenic zebrafish lines. S.H.O., P.S.C. and D.M.T. supervised the project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.M.T. (david.tobin@duke.edu).

METHODS

Zebrafish handling. All zebrafish husbandry and experimental procedures were performed in accordance and compliance with policies approved by the Duke University Institutional Animal Care and Use Committee (protocol A180-11-07). Clutches of eggs were collected from natural spawning and raised in filtered fish system water at 28 °C. Pigment development was halted in 1 dpf embryos by the addition of 1-phenyl-2-thiourea (PTU; Sigma-Aldrich; final concentration 45 µg ml⁻¹). Unless otherwise indicated, all zebrafish are from the wild-type AB strain.

Infection by microinjection. Embryos were anaesthetized at 2 dpf (or 1 dpf for morphants) with tricaine (MS-222; Sigma-Aldrich; final concentration 160 µg ml⁻¹) and injected with approximately 200 c.f.u. *M. marinum* or *E. coli* in an injection bolus of 10–20 nl as shown in Fig. 1b. Infected embryos were then recovered back to filtered fish system water supplemented with PTU. Embryos that were physically damaged by injection handling were discarded and excluded from further analysis.

Live imaging. Conventional and time-lapse fluorescent microscopy was carried out on a Zeiss Observer Z1 inverted microscope. Embryos were anaesthetized with tricaine and mounted in 3% (w/v) methylcellulose for static microscopy. Embryos for time-lapse microscopy were anaesthetized with 120 µg ml⁻¹ tricaine, mounted in 0.75% low melting point agarose in 96-well plates and immersed in filtered fish system water supplemented with PTU. Confocal microscopy was performed with an Olympus FV1000 confocal microscope. Images were processed with ImageJ (NIH), Photoshop CS4 (Adobe) and Velocity 5.4 image analysis software (Improvision/PerkinElmer Life and Analytical Sciences).

Image analysis. Abnormal vascular length was measured as the two-dimensional length of vessels not seen in control embryos in Photoshop using the ruler tool. Infection burden and neutrophil units were measured as the number of pixels above background per embryo in ImageJ using binary thresholding of single channel images and the Analyze Particles function. Macrophage association was scored as having *mpeg1:tdTomato-caax* expression around sites of Wasabi-fluorescent *M. marinum* expression. Dissemination was scored by comparing images of infected larvae at 4 and 6 dpi to track *M. marinum* infection foci.

Whole-mount *in situ* hybridization. Whole-mount *in situ* hybridization was carried out essentially as described³¹. Primers used for cloning *vegfaa* were previously described³². Primers used for cloning *phd3* were: 5'-ATTCCTGTGGGCTTCTCAAC-3' and 5'-ACACGAACCAACTGCTCAC-3'. Images of stained embryos were collected on a Nikon AZ100 microscope.

Construction of *mpeg1:tdTomato-caax* transgene. The previously characterized *mpeg1* promoter³³ was PCR amplified using the primers 5'-CCCAAACCTGAGTGTGTTGGAGCACATCTGACAT-3' and 5'-GGGAGGAAGCTTTGTTTGTCTCTGCACT-3'. The product was subsequently cloned into p5E MCS³⁴ using XhoI and HindIII sites to generate p5E *mpeg1*.

To construct a gateway-compatible *tdTomato-CAAX* construct, the sequence of *tdTomato* was PCR amplified with the primers 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTGGACCATGGTGAGCAAGGGCGAGGAG-3' and 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTAGATCTACTGTAGAGCTCGTCCATGCCG-3'. These primers introduced a silent SacI site in the 3' end of the *tdTomato* coding sequence and a BglII site downstream of the *tdTomato* stop codon. The PCR product was subsequently cloned by Gateway recombination (Invitrogen) into pDONR221 to generate pME *tdTomato*.

To generate pME *tdTomato-CAAX*, pME *Tomato* was digested with SacI and BglII and a linker sequence encoding the human H-Ras prenylation signal was cloned into the plasmid to generate pME *tdTomato-CAAX*. Primer sequences for the linker were as follows: top strand, 5'-CTACAAGAAGCTGAACCCCTCTGATGAGGTGGCCCGGCTGCATGAGCTGCAAGTGTGTGCTCTCTCA-3'; bottom strand, 5'-GATCTAGGAGAGCACACACTTGCAGCTCATGCAGCCGGGGC CACTCTCATCAGGAGGGTTTCAGCTTCTTGTAGAGCT-3'.

The *mpeg1:tdTomato-caax* transgene construct was subsequently constructed by recombining p5E *mpeg1*, pME *tdTomato-CAAX* and p3E polyA into pDestTol2pA2, by Gateway multisite recombination (Invitrogen) as previously characterized³⁴, to generate pDestTol2; *mpeg1:tdTomato-caax*.

Construction of *mfap4:turquoise* transgene. The *mfap4* promoter was PCR amplified using the primers 5'-CATGTTCTCGAGGCGTTTCTTGGTACAGCTGG-3' and 5'-CATGTTGGATCCACGATCTAAAGTCATGAAGAAAGA-3'. The product was subsequently cloned into p5E MCS using XhoI and BamHI sites. The native start codon was mutated using the primer 5'-CTGAGCTGTTGAGGAGAGAGTGAGAAG(ATT)GCAGTAAGTCTGTGGCTGTTTATTCC-3' by inverse PCR with the backbone primers 5'-GTAAGTCTGTGGCTGTTTATTTC-3' and 5'-CTTCTACTCTCTCTCAACAG-3'. The final p5E *mfap4* was then assembled by Gibson assembly using this single-stranded oligonucleotide and the backbone.

To generate pME *Turquoise2*, we used the primers 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTGGACCATGGTGAGCAAGGGCGAGGAG-3' and 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTTTACTTGTACAGCTCGTCC

AT-3' to amplify off pmTurquoise2 H2A (Addgene plasmid #36202; ref. 35). The PCR product was subsequently cloned into pDONR221 by BP cloning (Invitrogen) to generate pME *Turquoise2*.

The *mfap4:turquoise* transgene construct was subsequently constructed by recombining p5E *mfap4*, pME *Turquoise* and p3E polyA into pDestTol2pA2 to generate pDestTol2; *mfap4:turquoise*.

Construction of *lyzC:ntr-p2A-lanYFP* transgene. The previously characterized *lyzC* promoter was PCR amplified from the bacterial artificial chromosome (BAC) CH211-250A24 (BACPAC resources, Children's Hospital of Oakland Research Institute) using the primers 5'-CCCATAGGTACCCTGATCACTGGTGTAGTGAACCTC-3' and 5'-CCCAAACCTCGAGATTGTATCACTGCTGATATCTGCTTT-3'. The product was subsequently cloned into p5E MCS using KpnI and XhoI sites to generate p5E *lyzC*.

To generate pME *Ntr*, we used the primers 5'-GGGGACAAGTTTGTACAAA AAAAGCAGGCTGGACCATGGCCTCCGGACTCAGATCTCGAGC-3' and 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTCCACTTCGGTTAAGGTGATGTTTTGC-3' to amplify off pminiTol2-YFP-NTR (gift from L. Ramakrishnan). The PCR product was subsequently cloned into pDONR221 by BP cloning (Invitrogen) to generate pME *Ntr*.

To generate p3E p2A-lanYFP, we used the primers 5'-GGGGACAGCTTTC TTGTACAAAAGTGGTTGGATCCTTCAGTCTCGAGATGGTGAGCAAGGGC GAGGAG-3' and 5'-GGGGACAACCTTTGTATAATAAAGTTGTTACTTGTAC AGCTCGTCCAT-3' to amplify the dimeric form of lanYFP from pNCS-dlanYFP (Allele Biotechnology). The PCR product was subsequently cloned into pDONR P2R-P3 by BP cloning (Invitrogen) to generate p3E lanYFP. p3E lanYFP was subsequently digested with BamHI and XhoI and ligated with annealed oligonucleotides encoding the p2A sequence (GATCcGGAAGCGGAGCTACTAAGCTTCAG CTTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCCTGGACCTC and TCGAGAGTCCAGGGTTCCTCCACGCTCCAGCTCCAGCTGCTTCAGCAGG CTGAAGTTAGTAGTCCGCTTCCG³⁶) to generate p3E p2A lanYFP.

The *lyzC:ntr-p2A-lanYFP* transgene construct was subsequently constructed by recombining p5E *lyzC*, pME *Ntr* and p3E p2A-lanYFP into pDestTol2pA2 to generate pDestTol2; *lyzC:ntr-p2A-lanYFP*.

Nucleic acid microinjection. Tol2 transposase was generated from T3TS-Tol2 (ref. 37) using the mMessage mMachine T3 kit (Invitrogen). To generate transgenic zebrafish, embryos at the one-cell stage were injected with approximately 1 nl of transgenesis mixture consisting of 25 ng µl⁻¹ transposase RNA and 50 ng µl⁻¹ pDestTol2; transgenesis construct. Positive embryos were selected by fluorescence microscopy, raised to adulthood and transgenic founders were subsequently identified.

The Pu.1/Sp1 morpholino sequence was 5'-GATATACTGATACCTGATTGTGGT-3' and the Lt4h morpholino sequence was 5'-AGTAGGGTCTGAAAC TGGAGTCAT-3'. Morpholinos were injected at 10–60 µM.

Drug treatments. Metronidazole (M1547; Sigma-Aldrich; final concentration 5 mM) was dissolved in water; the relatively high concentration of metronidazole seems to be a function of bioavailability in zebrafish as higher concentrations are routinely required for nitroreductase-mediated cellular ablation studies³⁸. Pazopanib (sc-364564; Santa Cruz Biotechnology; final concentration 250 nM for larvae, 1 µM for adults) and SU5416 (S8442; Sigma-Aldrich; final concentration 250 nM) were dissolved in DMSO. Metronidazole, pazopanib and SU5416 were added immediately after infection and refreshed every 2 days for the duration of the experiment. Randomization into drug treatment groups was achieved by random selection of infected zebrafish from a single pool before addition of drugs.

Microangiography. To detect vascular leakiness in wild-type and *Tg(flk1:eGFP)* embryos, embryos were anaesthetized in tricaine water and injected with a 10 nl bolus of dextran-Texas Red 70,000 MW (D-1830; Life Technologies; 1 mg ml⁻¹ final concentration) into the posterior section of the dorsal aorta or posterior cardinal vein. This injection location avoided injection-trauma-induced tissue leakage occurring near *M. marinum* lesions. Injected embryos were rinsed in tricaine water and immediately mounted in methylcellulose for fluorescent microscopy. Vascular leakage was calculated as a ratio of intersomitic dextran-Texas Red signal divided by aortic dextran-Texas Red signal.

Combined drug treatments. Rifampicin (R3501; Sigma-Aldrich; final concentration 50 µM) was dissolved in DMSO. To achieve a suboptimal dose of VEGFR inhibition, SU5416 was added immediately after infection at a final concentration of 200 nM and refreshed at 3 dpi. Rifampicin was added at 3 dpi and not refreshed for the duration of the experiment.

Metronidazole (final concentration 5 mM) and pazopanib (suboptimal final concentration 200 nM) were added immediately after infection and refreshed every 2 days for the duration of the experiment.

Randomization into drug treatment groups was achieved by random selection of infected embryos from a single pool prior to addition of drugs; blinding was not performed for subsequent quantitation.

Adult infections. Adult zebrafish were infected with approximately 500 c.f.u. of fluorescent *M. marinum* via intraperitoneal injection. Zebrafish were maintained in beakers in a dedicated incubator at 28 °C with a 14:10 h light:dark cycle. Pazo-panib was added to a final concentration of 1 µM immediately after infection and refreshed every 2 days for the duration of the experiment.

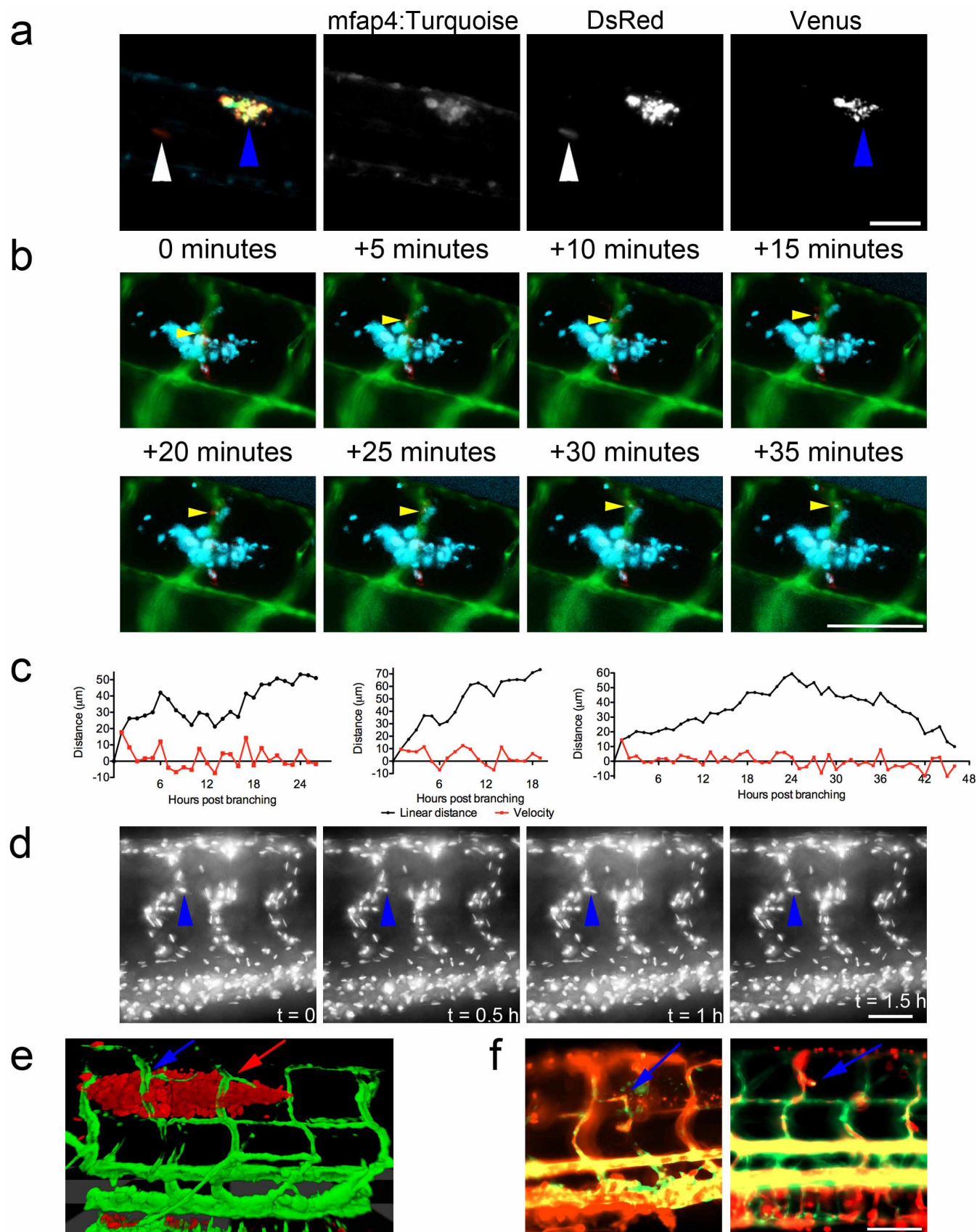
Randomization into drug treatment groups was achieved by random selection of infected fish from a single pool prior to addition of drugs; blinding was not performed for subsequent quantitation.

Bacterial recovery from infected adults. Infected adult zebrafish were pre-treated with 25 µg ml⁻¹ hygromycin to reduce microbiota load for 2 h before harvesting. Zebrafish were euthanized by tricaine overdose and homogenized by bead mill for three bursts of 15 s. Homogenate was plated on Middlebrook 7H10 (262710; Difco) supplemented with OADC, hygromycin (H0654; Sigma-Aldrich, 50 µg l⁻¹) and amphotericin B (SV3007801; Thermo Scientific, 10 mg l⁻¹). Plates were grown at 30 °C for 10–14 days until fluorescent colonies could be counted.

Hypoxyprom staining. Infected adult zebrafish were injected with 15 µl of a 10 mg ml⁻¹ pimonidazole solution (HP7; Hypoxyprom) every 2 days from 8 dpi to 14 dpi. Zebrafish were euthanized by tricaine overdose at 14 dpi, fixed in 4% PFA, decalcified in 0.5 M EDTA and cyrosectioned. Frozen sections were stained with 4.3.1.3 mouse Dylight 549-MAb (HP7; Hypoxyprom) or with unconjugated 4.3.1.3 mouse monoclonal antibody and secondary detection was carried out with goat anti-mouse Alexa-Fluor 647 (A-21235; Life Technologies) to detect hypoxic cells.

Statistics. Data are presented as mean ± s.d. Experiments were analysed with the statistical tests indicated in figure legends using Prism 5 (Graphpad). Unpaired Student's *t*-tests were performed unless otherwise indicated. For ANOVA analyses with Tukey's post-test, *P* values are indicated as follows: **P* < 0.05, ***P* < 0.01, ****P* < 0.001.

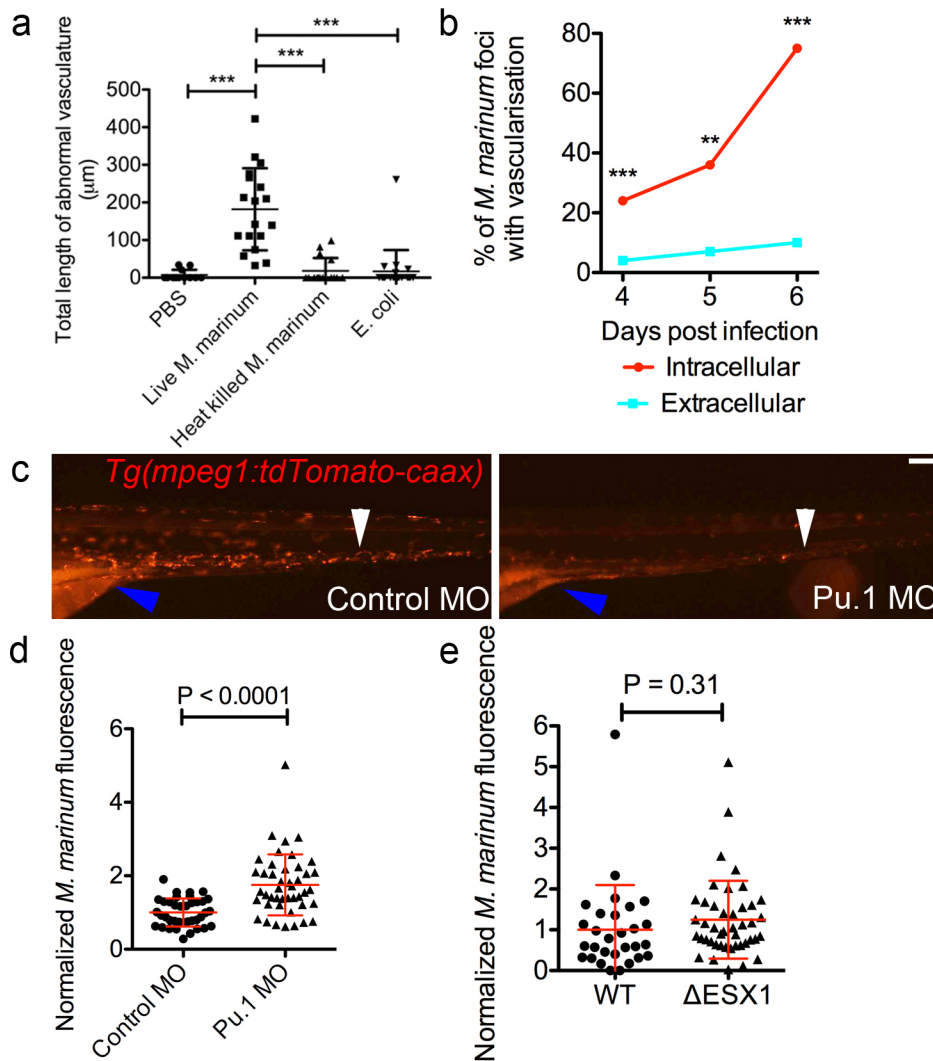
31. Thisse, C. & Thisse, B. High-resolution *in situ* hybridization to whole-mount zebrafish embryos. *Nature Protocols* **3**, 59–69 (2008).
32. Liang, D. *et al.* Cloning and characterization of vascular endothelial growth factor (VEGF) from zebrafish, *Danio rerio*. *Biochim. Biophys. Acta* **1397**, 14–20 (1998).
33. Ellett, F., Pase, L., Hayman, J. W., Andrianopoulos, A. & Lieschke, G. J. *mpeg1* promoter transgenes direct macrophage-lineage expression in zebrafish. *Blood* **117**, e49–e56 (2011).
34. Kwan, K. M. *et al.* The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.* **236**, 3088–3099 (2007).
35. Goedhart, J. *et al.* Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nature Commun.* **3**, 751 (2012).
36. Kim, J. H. *et al.* High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS ONE* **6**, e18556 (2011).
37. Balciunas, D. *et al.* Harnessing a high cargo-capacity transposon for genetic applications in vertebrates. *PLoS Genet.* **2**, e169 (2006).
38. Curado, S., Stainier, D. Y. & Anderson, R. M. Nitroreductase-mediated cell/tissue ablation in zebrafish: a spatially and temporally controlled ablation method with applications in developmental and regeneration studies. *Nature Protocols* **3**, 948–954 (2008).



Extended Data Figure 1 | Angiogenesis in the zebrafish *M. marinum*

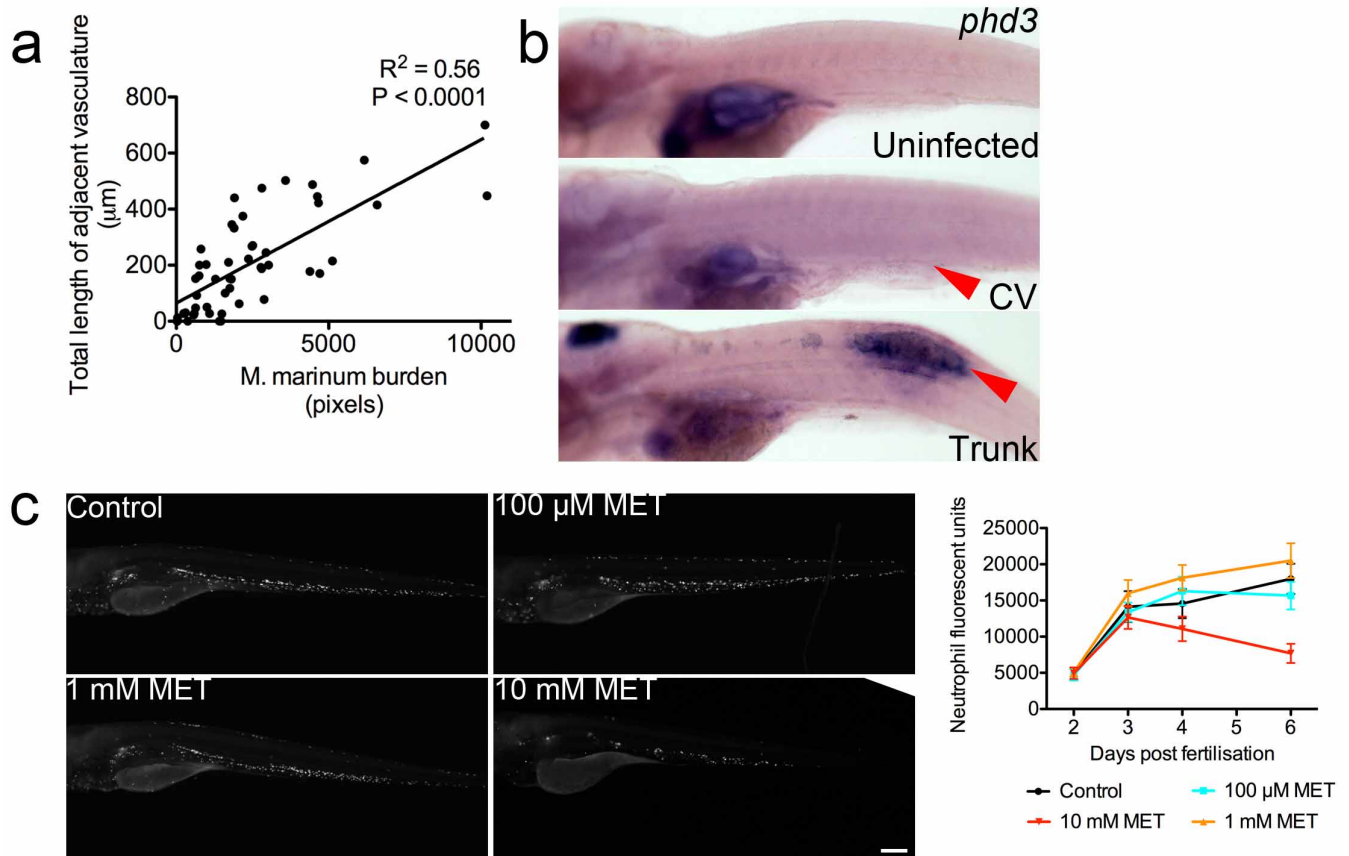
infection model. **a**, Image of 6 dpi *Tg(mfap4:turquoise^{xt27})* larvae infected with *M. marinum* SM2 pMAP49::Venus. Blue arrowheads indicates site of granuloma with induced expression of Venus from phagocytosed *M. marinum*. White arrowheads indicate sites of extracellular *M. marinum* growth detected by constitutive DsRed expression but no macrophage-induced Venus expression. Image is representative of granulomas found in five individual animals. **b**, Time-lapse images of Cerulean-fluorescent *M. marinum* dissemination from an established granuloma into the adjacent intersegmental vessel in a *Tg(flk1:eGFP, mpeg1:tdTomato-caax^{xt3})* double-transgenic larva where bacterial are labelled blue, blood vessels are labelled green and macrophages are labelled red. Yellow arrow tracks a single infected macrophage egressing the established granuloma and entering the vasculature. Images are representative of macrophage behaviour in three individual animals. **c**, Plots of

vessel growth kinetics from three individual branches in individual *Tg(flk1:eGFP)* larvae. Videos of each larva analysed are available in Supplementary Videos 6 and 7 (left), and 8 and 9 (right). **d**, Time-lapse images of nuclear division during vascular growth in a single *Tg(fli1a:eGFP-nls)* larva. Blue arrowhead indicates nucleus of interest. Images are representative of nuclear division in ten individual animals. Video of nuclear division is available in Supplementary Video 10. **e**, Three-dimensional rendering of recruited blood vessels in a *Tg(flk1:eGFP)* larva infected with Tomato-fluorescent *M. marinum* originating from arterial and venous ISVs as indicated by red and blue arrows, respectively. Image is representative of ten individual animals. **f**, Extended exposure images of blood flow in *Tg(flk1:eGFP, gata1:DsRed^{sd2})* larvae. Blue arrows indicate blood flow through ectopic vessels. Images are representative of blood flow in 20 individual animals. Scale bars, 100 μ m.



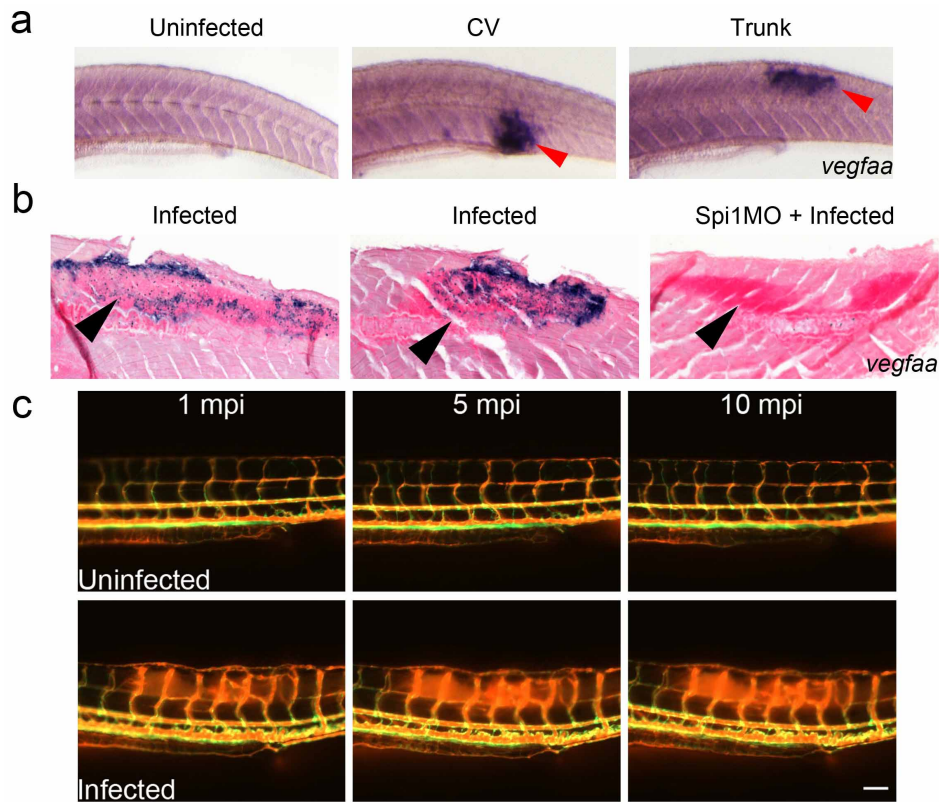
Extended Data Figure 2 | Formation of ectopic vasculature is dependent on granuloma formation. **a**, Length of abnormal vasculature in *Tg(flk1:eGFP)* larvae injected with PBS, live *M. marinum*, heat-killed *M. marinum* and *E. coli*. One-way ANOVA with Tukey's post-test, data are representative of two biological replicates. **b**, Recruitment of vasculature by intracellular and extracellular foci of *M. marinum*. Total number of foci analysed: 4 dpi, 221 intracellular, 105 extracellular; 5 dpi, 71 intracellular, 26 extracellular; and 6 dpi, 131 intracellular, 50 extracellular. Fisher's exact test. **c**, Comparative images of 5 dpf control and Pu.1 morphant *Tg(mpeg1:tdTomato-caax^{xt3})* larvae. White

arrowhead indicates comparative locations within the caudal haematopoietic tissue. Blue arrowhead indicates intestinal and yolk sac autofluorescence. Scale bar, 100 μm. Images are representative of transgene expression in 20 animals per treatment group. **d**, **e**, Bacterial burden in 5 dpi control and Pu.1 morphant larvae (**d**), and 4 dpi larvae infected with wild-type (WT) or ΔESX1 Tomato-fluorescent *M. marinum* (**e**). Student's *t*-test with Welch's correction, all data are pooled from two biological replicates. Error bars represent mean ± s.d. ***P* < 0.01, ****P* < 0.001.



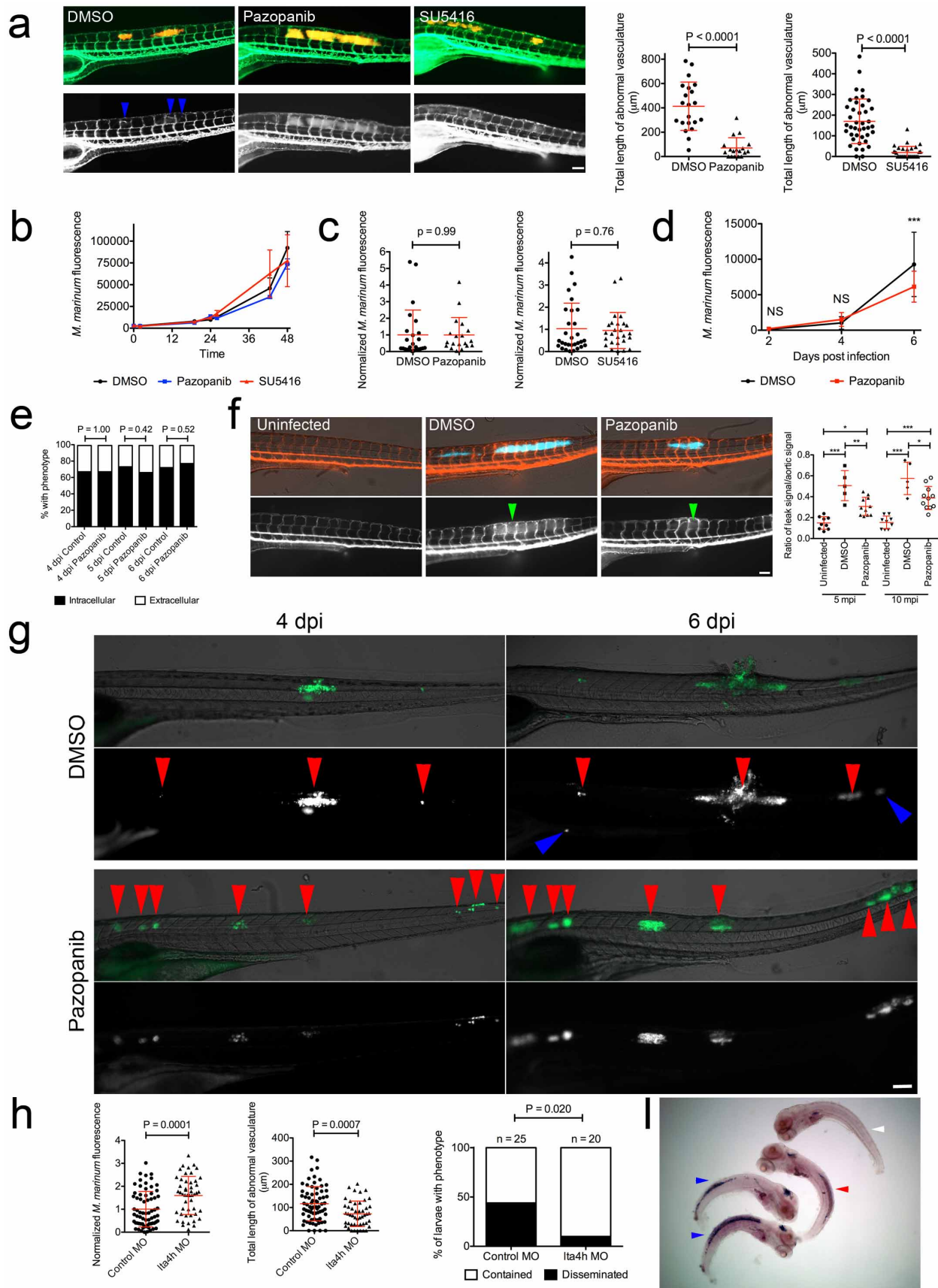
Extended Data Figure 3 | Granuloma vascularization correlates with granuloma size. **a**, Plot of abnormal vasculature length and bacterial burden for individual foci of infection measured by fluorescent pixel count (FPC) in *Tg(flk1:eGFP)* larvae. Slope significantly not zero, $P < 0.0001$ linear regression, data are pooled from three biological replicates. **b**, Whole-mount *in situ* hybridization detection of *phd3* expression. Images are representative of *phd3* staining in uninfected (20/20), caudal vein (CV)-infected (20/20) and

trunk-infected (7/20) zebrafish. **c**, Left, images of *Tg(lyzC:ntr-p2A-lanYFP^{xt14})* larvae treated with metronidazole as indicated. Green arrowheads indicate comparative locations within caudal haematopoietic tissue. Images are median images from experimental groups: control, $n = 21$; 100 μM , $n = 22$; 1 mM, $n = 24$; and 10 mM, $n = 19$. Right, quantification of neutrophil numbers by area of fluorescence in *Tg(lyzC:ntr-p2A-lanYFP^{xt14})* larvae treated with metronidazole from 2 dpf to 6 dpf. Error bars represent mean \pm s.d.



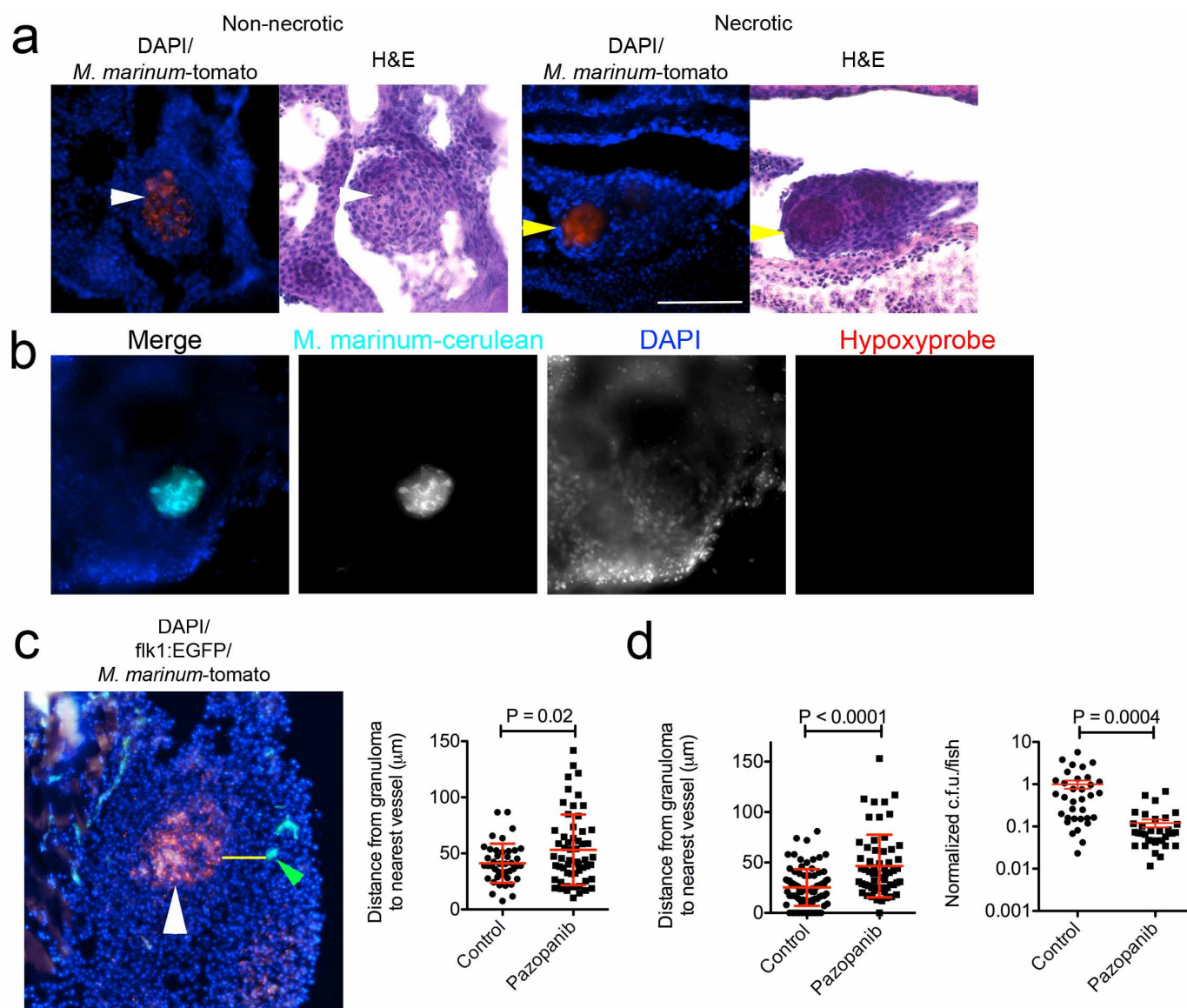
Extended Data Figure 4 | *M. marinum* infection induces expression of *vegfaa*. **a**, Whole-mount *in situ* hybridization detection of *vegfaa* expression in uninfected, caudal vein (CV)-injected and trunk-injected larvae. Red arrow indicates sites of infection with *vegfaa* expression. Images are representative of 20 animals per treatment group. **b**, Representative histological sections of whole-mount *in situ* hybridization detected *vegfaa* expression in control infected larvae and a Pu.1 morpholino (MO)-treated infected larva. Black

arrows indicate sites of infection identified by increased nuclear fast red staining density. Images are representative of ten animals per treatment group. **c**, Microangiography of *Tg(flk1:eGFP)* larvae imaged at 1, 5 and 10 min post-injection (mpi). Top panels are representative of uninfected larvae, bottom panels are representative of larvae infected with unlabelled *M. marinum*. Images are representative of ten animals per treatment group. Scale bars, 100 μ m.



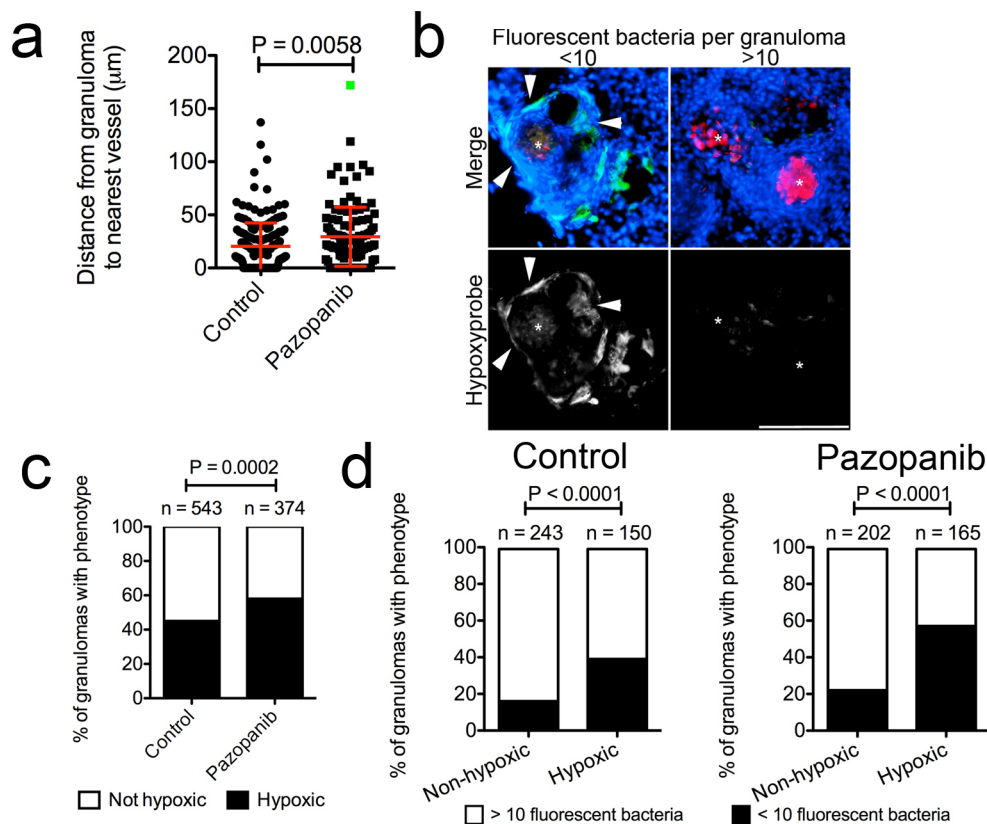
Extended Data Figure 5 | Pazopanib and SU5416 reduce *M. marinum* pathogenicity in zebrafish larvae. **a**, Left, comparative images of *Tg(flk1:eGFP)* larvae infected with Tomato-fluorescent *M. marinum* and treated with DMSO, pazopanib or SU5416. Top panels depict Tomato-fluorescent *M. marinum* and labelled vasculature. Bottom panels depict only *Tg(flk1:eGFP)*-labelled vasculature. Blue arrowheads indicate somites with ectopic vasculature. Images are representative of 20 animals per treatment group. Right, length of abnormal vasculature in pazopanib- or SU5416-treated larvae. Student's *t*-test, data are pooled from two or three biological replicates, respectively. **b**, Growth curve of Tomato-fluorescent *M. marinum* in 7H9 broth culture supplemented with pazopanib or SU5416. Data are representative of two biological replicates. **c**, Bacterial burden in caudal-vein-infected larvae treated with either pazopanib or SU5416. Student's *t*-test, data are pooled from two biological replicates. **d**, Longitudinal bacterial burden from 2 to 6 dpi in trunk-infected larvae treated with pazopanib. One-way ANOVA with Tukey's post-test. NS, not significant; $n=14$ individuals per group. **e**, Comparison of *M. marinum* foci between control and pazopanib-treated larvae scored by association with macrophages. Fisher's exact test, $n=40$ individuals per group. **f**, Left, microangiography of larvae infected with cerulean-fluorescent *M. marinum*, injected with high-molecular-weight

dextran-Texas Red at 6 dpi and imaged at 5 minutes post dextran injection (mpi). Top panels depict Cerulean-fluorescent *M. marinum* and dextran-Texas Red, bottom panels depict only dextran-Texas Red in vasculature and leakage around sites of infection. Green arrowheads indicate somites with the highest leakage signals in infected larvae. Images are median images from graph on right. Right, quantification of vascular leakage in uninfected, DMSO- and pazopanib-treated larvae. One-way ANOVA with Tukey's post-test, data are representative of two biological replicates. **g**, Dissemination of Wasabi-fluorescent *M. marinum* in larvae treated with DMSO or pazopanib. Red arrowheads indicate contained foci of infection that remain in the same location throughout the course of infection, blue arrowheads indicate disseminated foci of infection. Images are representative of data in Fig. 3b. **h**, Bacterial burden (left), length of abnormal vasculature (middle) and dissemination (right) in 5 dpi control and *Lta4h* morphant larvae. **i**, Whole-mount *in situ* hybridization detection of *phd3* expression in uninfected (white arrow) and *M. marinum*-infected zebrafish larvae. Blue arrows indicate *phd3*-expression-positive larvae with purple staining, red arrow indicates site of bacterial infection with no purple staining, indicating *phd3*-expression-negative larva. Image is representative of data in Fig. 3c. Scale bars, 100 μm . Error bars represent mean \pm s.d. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.



Extended Data Figure 6 | Effects of pazopanib treatment are reproduced in adult zebrafish infections. **a**, Images of non-necrotic (left) and necrotic (right) Tomato-fluorescent *M. marinum* granulomas stained with DAPI (top) and haematoxylin and eosin (bottom). White arrows indicate non-necrotic granuloma, yellow arrows indicate necrotic granuloma. Images are representative of granulomas found in eight individual animals. **b**, Representative image of a necrotic granuloma from a negative control, not injected with pimnidazole, 2 wpi adult *Tg(flk1:eGFP)* zebrafish infected with cerulean-fluorescent *M. marinum* (cyan), and stained for hypoxypore (red) and with DAPI (blue). Images are representative of granulomas found in two individual animals. **c**, Left, representative image of Tomato-fluorescent *M. marinum* granuloma in *Tg(flk1:eGFP)* zebrafish stained with DAPI. White

arrow indicates granuloma, yellow line indicates path measured for distance between granuloma and nearest vasculature (indicated by green arrow). Image is representative of data presented on the right, in panel **d** and Extended Data Fig. 7a. Right, distance between granulomas and nearest vasculature measured in 2 wpi adult *Tg(flk1:eGFP)* zebrafish. Total number of zebrafish analysed: 4 (control), 4 (pazopanib). **d**, Left, distance between granulomas and nearest vasculature measured in 2 wpi adult *Tg(flk1:eGFP)* zebrafish treated with pazopanib for 1 week. Total number of zebrafish analysed: 2 (control), 2 (pazopanib). Right, bacterial burden in 2 wpi adult zebrafish treated with pazopanib for 1 week. Student's *t*-test, data are pooled from three biological replicates.



Extended Data Figure 7 | Pazopanib increases the frequency of hypoxic and low-burden granulomas. **a**, Distance between granulomas and nearest vasculature measured in 6 wpi adult *Tg(flk1:eGFP)* zebrafish. Total number of zebrafish analysed: 4 (control), 4 (pazopanib). Green dot indicates outlier that was omitted from statistical analysis. **b**, Images of low burden/hypoxic (left) and high burden/non-hypoxic (right) granulomas in zebrafish that were injected with pimonidazole. Asterisks indicate Tomato-fluorescent *M. marinum*, arrows indicate areas of hypoxia in granuloma. Images are

representative of data in **c**, **d** and Fig. 4d. **c**, Comparison of granulomas between control and pazopanib-treated adult zebrafish scored for pimonidazole staining. Total number of zebrafish analysed: 4 (control), 4 (pazopanib). **d**, Comparison of granulomas between non-hypoxic and hypoxic granulomas in control and pazopanib-treated adult zebrafish scored for *M. marinum* burden. Total number of zebrafish analysed: 4 (control), 4 (pazopanib). Scale bars, 100 μm . Error bars represent mean \pm s.d.

p63⁺Krt5⁺ distal airway stem cells are essential for lung regeneration

Wei Zuo¹, Ting Zhang¹, Daniel Zheng¹, An Wu¹, Shou Ping Guan¹, Audrey-Ann Liew¹, Yusuke Yamamoto², Xia Wang², Siew Joo Lim¹, Matthew Vincent³, Mark Lessard⁴, Christopher P. Crum⁵, Wa Xian^{1,2,5,6,7} & Frank McKeon^{1,2,6}

Lung diseases such as chronic obstructive pulmonary disease¹ and pulmonary fibrosis² involve the progressive and inexorable destruction of oxygen exchange surfaces and airways, and have emerged as a leading cause of death worldwide. Mitigating therapies, aside from impractical organ transplantation, remain limited and the possibility of regenerative medicine has lacked empirical support. However, it is clinically known that patients who survive sudden, massive loss of lung tissue from necrotizing pneumonia^{3,4} or acute respiratory distress syndrome^{5,6} often recover full pulmonary function within six months. Correspondingly, we recently demonstrated lung regeneration in mice following H1N1 influenza virus infection, and linked distal airway stem cells expressing Trp63 (p63) and keratin 5, called DASC^{p63/Krt5}, to this process⁷. Here we show that pre-existing, intrinsically committed DASC^{p63/Krt5} undergo a proliferative expansion in response to influenza-induced lung damage, and assemble into nascent alveoli at sites of interstitial lung inflammation. We also show that the selective ablation of DASC^{p63/Krt5} *in vivo* prevents this regeneration, leading to pre-fibrotic lesions and deficient oxygen exchange. Finally, we demonstrate that single DASC^{p63/Krt5}-derived pedigrees differentiate to type I and type II pneumocytes as well as bronchiolar secretory cells following transplantation to infected lung and also minimize the structural consequences of endogenous stem cell loss on this process. The ability to propagate these cells in culture while maintaining their intrinsic lineage commitment suggests their potential in stem cell-based therapies for acute and chronic lung diseases.

H1N1 influenza virus infection of murine lung triggers a process of leukocyte infiltration and lung damage similar to that of acute respiratory distress syndrome (ARDS)^{7–10} (Fig. 1a). Damaged regions are marked by densely packed, CD45⁺ neutrophils and macrophages¹¹ and an absence of markers for type I (Pdpn⁺) and type II (SPC⁺) pneumocytes (Fig. 1a; Extended Data Fig. 1a). Despite the local destruction of alveoli, these same regions harbour discrete clusters of p63⁺Krt5⁺ epithelial cells proposed to be the early stages of *de novo* alveoli formation (Fig. 1a)⁷. Three dimensional reconstruction of serial sections of lungs at 15 days post-infection (dpi) reveal a broad distribution of Krt5⁺ cells along the axis of the bronchioles (Fig. 1b; Supplementary Video 1).

To decipher the origin of these p63⁺Krt5⁺ cells appearing in response to lung damage, we performed genetic lineage-tracing of Krt5⁺ cells starting before infection through the cycle of lung damage and resolution. Mice expressing a tamoxifen-dependent lacZ gene under the control of the Krt5 promoter¹² (Tg (KRT5-Cre^{ERT2}) ROSA26-lsl-lacZ)¹³ were treated with tamoxifen before intratracheal delivery of H1N1 influenza virus (Fig. 1c; Extended Data Fig. 1b). At 0 dpi, lacZ activity was not detectable in whole-mount lung and yet Krt5⁺ cells were evident in distal lung as clusters of peribronchiolar Krt5⁺p63⁺ cells were observed in one of three to four consecutive sections of lung. Approximately 50% of these cells expressed *Escherichia coli*-specific β -galactosidase (7 clusters of p63⁺ cells were observed in 25 slides with a total of 39 p63⁺ cells, of

which 19 were β -galactosidase positive; Fig. 1c). No labelling of other cell types in the lung was observed. In addition, colonies of distal airway stem cells (DASC) with long-term self-renewal (passage 4) were generated^{7,14,15} from three 0 dpi mice and stained with antibodies to Krt5

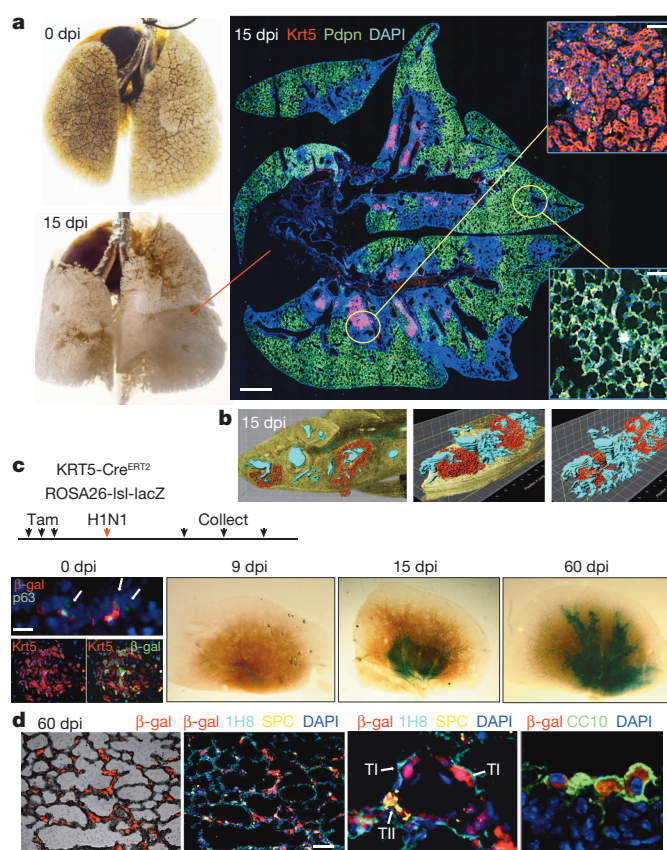


Figure 1 | Lineage tracing of Krt5⁺ cells following viral infection. **a**, Left, mouse lung before and after viral infection. Right, immunofluorescence images of infected lung of anti-Krt5 (red), anti-Pdpn (green) with DNA counterstain (DAPI, blue). Scale bar, 1 mm. Insets, high magnification of indicated regions. $n = 10$ mice. Scale bars, 100 μ m. **b**, Three-dimensional reconstruction of anti-Krt5 (red) from serial sections of infected lung (bronchioles, blue). Grid, 100 \times 100 μ m. **c**, Schematic of lineage tracing experiment following tamoxifen treatment to reveal lacZ expression 0, 9, 15 and 60 days post infection ($n = 3$ mice for each time point and control). At 0 dpi, rare clusters of p63⁺ cells in bronchioles express *E. coli* β -galactosidase (left top, arrows). Scale bar, 20 μ m. Colony of Krt5⁺ DASC co-expressing β -galactosidase at 0 dpi (left bottom). At 9, 15 and 60 dpi, whole lungs stained by X-gal. **d**, Immunofluorescence in 60 dpi lung sections with indicated antibodies. TI, TII indicate type I and II pneumocytes, respectively. Scale bar, 100 μ m.

¹Genome Institute of Singapore, A-STAR, 138672 Singapore. ²The Jackson Laboratory for Genomic Medicine, Farmington, Connecticut 06032, USA. ³Advanced Cell Technologies, Marlborough, Massachusetts 01752, USA. ⁴The Jackson Laboratory, Bar Harbor, Maine 04609, USA. ⁵Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁶Department of Medicine, National University Health System, 119228 Singapore. ⁷Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmington, Connecticut 06030, USA.

and β -galactosidase. All 256 colonies stained with anti-Krt5 antibodies and 132 co-expressed β -galactosidase (Fig. 1c; Extended data Fig. 1c). At 9 dpi, lacZ activity was subtle and restricted to the airways (Fig. 1c), consistent with our previous observations that at 9 dpi, $p63^{+}Krt5^{+}$ cells had accumulated within mouse bronchioles with no evidence of migration to interstitial regions⁷. By 15 dpi, however, the lacZ signal was significantly more robust and included broader regions of interstitial lung (Fig. 1c; Extended Data Fig. 1d). At 60 dpi, the lacZ signal was distributed along the conducting airways and the surrounding interstitial regions, suggesting a progressive process (Fig. 1c). Importantly, no lacZ activity was detected in the lungs of tamoxifen-treated mice in the absence of infection, indicating that the robust signal we observed was in response to lung damage (Extended data Fig. 1e). Histological analysis of the lacZ-positive regions of lung from infected mice revealed broad interstitial areas of staining corresponding to alveoli and that $72 \pm 7\%$ of 1,051 lineage-labelled cells expressed type I ($1H8^{+}$ and $Pdpn^{+}$) or type II (SPC^{+}) pneumocyte markers with the remainder being secretory cells in the bronchioles ($n = 3$ mice; Fig. 1d; Extended Data Fig. 1d, f).

To generate a mouse model in which $DASC^{p63/Krt5}$ could be conditionally ablated, we engineered the human diphtheria toxin receptor (DTR)¹⁶ into the *Krt6a* (*Krt6*) locus (Krt6-DTR; Fig. 2a; Extended Data Fig. 2a, b) as the *Krt6* gene becomes activated specifically in DASCs

at approximately 8–10 dpi⁷, rather than the *p63* or *Krt5* gene that are expressed in stem cells of many stratified epithelia¹⁷. Consistently, $DASC^{p63/Krt5}$ from the Krt6a-DTR mice were found to co-express DTR and Krt6a *in vivo* and *in vitro* (Extended Data Fig. 2c, d). To test this ablation model *in vivo*, we infected the Krt6-DTR mice with influenza virus and injected diphtheria toxin at 8 dpi (Fig. 2a). By 15 dpi, exposure to toxin resulted in a rapid loss of interstitial clusters of $Krt5^{+}Krt6^{+}$ cells (Fig. 2b, c), indicating that we had generated a highly efficient ablation model. Compared to wild-type controls, Krt6-DTR mice lost 90% of $Krt5^{+}$ cells and over 99% of $Krt6^{+}$ cells following diphtheria toxin treatment (Fig. 2c).

We next evaluated the effect of ablating $DASC^{p63/Krt5}$ on the process of lung regeneration itself. Focal infiltrates of leukocytes appearing as haematoxylin-eosin staining densities form at about 10–15 dpi in wild-type mice and are typically resolved over the next 15–45 days⁷ (Fig. 2d, e). However, in diphtheria-toxin-treated Krt6-DTR mice, these densities fail to resolve over time (Fig. 2d, e). The persistent damage in $DASC^{p63/Krt5}$ -ablated lungs is also evident by whole-lung genome expression analyses (Extended Data Fig. 3a, false discovery rate (FDR) $q < 0.001$). Histological comparisons of the persistent densities at 30 dpi revealed the possible basis for this difference. Although wild-type lungs at 30 dpi still display some lung densities, nearly all of these densities were negative

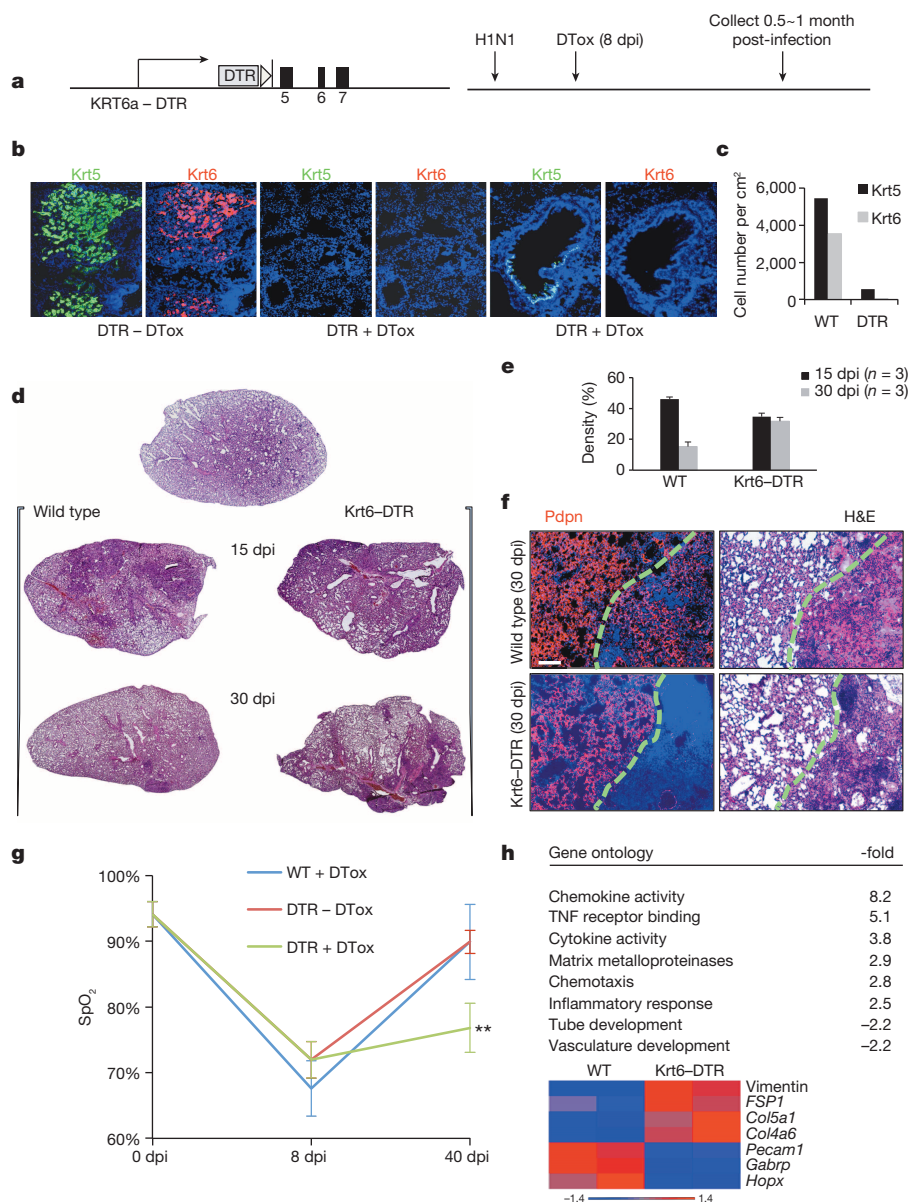


Figure 2 | Conditional ablation of activated $DASC^{p63/Krt5}$. **a**, Modified *Krt6a* locus driving the human diphtheria toxin receptor and experimental scheme. DTTox, diphtheria toxin. **b**, Immunofluorescence images of distal lung at 15 dpi with indicated diphtheria toxin (+/–DTTox) condition. **c**, Quantification of $Krt5^{+}$ and $Krt6^{+}$ cells in DTTox-treated mice. $n = 2$ mice per group, 6 sections covering the whole lung for each mouse. **d**, Lung sections from indicated mice following influenza virus infection and treatment with DTTox. **e**, Morphometric analysis of interstitial densities in sections of lung ($n = 3$ mice per condition and time). Error bars, s.e.m. **f**, Anti-Pdpn (red) immunofluorescence of lung densities to reveal type I pneumocytes counterstained with DAPI (blue). Scale bar, 100 μ m. H&E, haematoxylin and eosin. **g**, Peripheral capillary oxygen saturation (SpO₂) values obtained by pulse oximetry (WT, $n = 3$, DTR \pm DTTox, $n = 4$) at indicated times. Error bars, s.e.m. ** $P < 0.01$ for –DTTox versus +DTTox. **h**, Gene ontology classes and fold-change $DASC^{p63/Krt5}$ -ablation versus normal mouse lung at 30 dpi. Below, heat map of differentially expressed ($P < 0.05$) genes involved in pre-fibrosis, alveolar structure and vasculature.

for the CD45 leukocyte marker and instead possess unusual networks of Pdpn⁺ type I pneumocytes (Fig. 2f; Extended Data Fig. 3b). These networks of type I pneumocytes lacked markers of type II pneumocytes (Extended Data Fig. 4). Interestingly, these networks stain positive for other type I pneumocyte markers (Aqp5) but not all (for example, Hopx; Extended Data Fig. 4), suggesting the possibility that the type I pneumocytes in these networks are undergoing a maturation process. Remarkably, similar alveoli-like structures formed by type I pneumocytes lacking type II pneumocytes in anatomical analyses of mice recovering from infection by the NWS influenza A virus were reported nearly 40 years ago¹⁸. In contrast to the type I pneumocytes networks in persistent densities of wild-type mice, DASC^{p63/Krt5}-ablated mice showed no such alveolar networks at 30 dpi (Fig. 2f) and instead maintained persistent infiltration of leukocytes evidenced by anti-CD45 staining (Extended Data Fig. 3b).

We next asked if the loss of DASC^{p63/Krt5} also affected aspects of pulmonary function in these mice. Using pulse oximetry¹⁹ to assess peripheral capillary oxygen saturation (SpO₂), we found that the normal 95% SpO₂ values plummeted to approximately 70% in normal and DASC^{p63/Krt5}-ablated mice eight days following infection. However, the normal mice recovered to 90% SpO₂ by 40 dpi, whereas the DASC^{p63/Krt5}-ablated mice only reached SpO₂ values approaching 75% saturation (Fig. 2g). Consistent with this apparent decline in pulmonary function, the persistent densities in the 30 dpi Krt6-DTR lung showed staining for smooth muscle actin (α -SMA; Extended Data Fig. 5), a marker of myofibroblasts known to be associated with a pre-fibrotic state of the lung²⁰. These same interstitial regions showed weak but detectable staining with Masson's trichrome blue, a marker of fibrosis (Extended Data

Fig. 5). Correspondingly, whole-genome expression profiles of wild-type and DASC^{p63/Krt5}-ablated lungs indicated the persistence of inflammatory gene expression and a relative decrease in gene expression linked to vasculature development in the DASC^{p63/Krt5}-ablated lungs at 30 dpi (Fig. 2h; Extended Data Fig. 5). Moreover, the DASC^{p63/Krt5}-ablated mice showed the presence of pre-fibrosis gene signature^{21,22} including vimentin, FSP1 and collagen genes (Fig. 2h). Together these data suggest that the ablation of DASC^{p63/Krt5} arising during acute injury results in a failure of the regenerative process with structural and functional consequences for the lung.

Whereas p63⁺Krt5⁺ cells are prominent features of the proximal lung, their presence in distal lung has been less clear²³. This is reflected in the abundant Krt5⁺ cells in proximal lung (Fig. 3a) and the intermittent, peri-bronchiolar clusters seen in one of every three to four sections of distal lung (Fig. 3a, arrows). Consistently, 100-fold fewer DASC^{p63/Krt5} colonies arise from cell suspensions of distal lung than tracheobronchial stem cell (TBSC^{p63/Krt5}) colonies from proximal lung (Extended data Fig. 6a). Regardless, both DASC^{p63/Krt5} and TBSC^{p63/Krt5} can be cloned and propagated as single-cell-derived pedigrees⁷ that show very different fates upon differentiation. In air-liquid interface (ALI) cultures²⁴, TBSCs yield a stratified epithelium with Krt5⁺ basal cells and apical ciliated and secretory cells typical of proximal airway, while DASCs yield a monolayer of differentiated cells expressing Pdpn (Fig. 3b; Extended Data Fig. 6b). In three-dimensional Matrigel cultures, DASCs form unilaminar, alveolar-like spheres composed of cells expressing type I (Pdpn and Aqp5) and type II (SPC) pneumocyte markers (Fig. 3c). The very minor variation in gene expression between TBSCs and DASCs

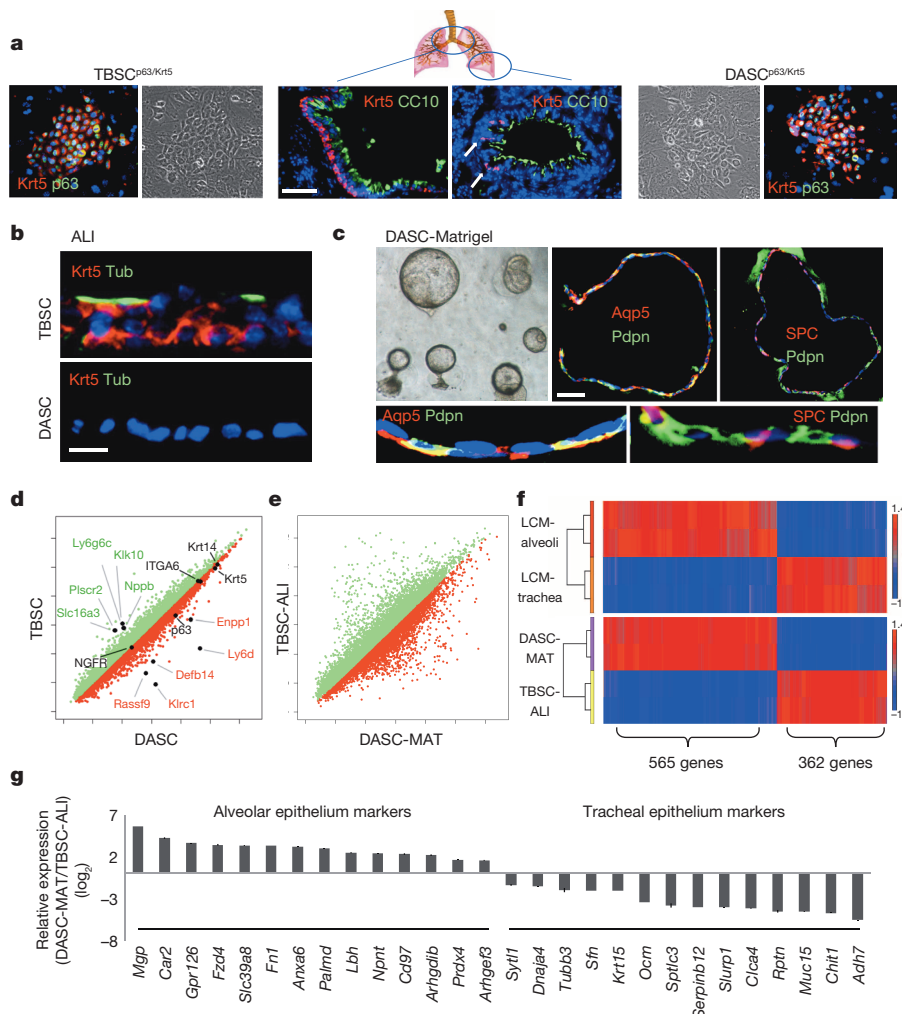


Figure 3 | Cloning and *in vitro* differentiation of DASCs and TBSCs. **a**, Krt5⁺ cells in proximal (left) and distal (arrows, right) lung, with corresponding TBSC and DASC colonies (outside panels). Scale bar, 50 μ m. **b**, Differentiation of TBSC and DASC in air-liquid interface cultures showing respectively stratified epithelia with ciliated (acetylated-Tub⁺) cells and a monolayer of differentiated cells. Scale bar, 20 μ m. **c**, Unilaminar spheres formed by DASCs in three-dimensional Matrigel cultures and the expression of indicated markers in sections. Scale bar, 50 μ m. Insets, high magnification. **d**, Scatter plot of gene expression of immature TBSCs and DASCs highlighting common and disparately expressed (fold change > 3, $P < 0.001$) genes. **e**, Scatter plot of gene expression of TBSC-ALI and DASC-MAT. **f**, Top, heat map of gene sets differentially expressed in murine alveoli and tracheal epithelium ($P < 0.05$). Bottom, heat map of differentially expressed genes in DASC-MAT versus TBSC-ALI ($P < 0.05$) informed by alveolar and tracheal data sets. **g**, Histogram of differentially expressed genes of DASC-MAT versus TBSC-ALI for which validating immunohistochemistry data are available (see <http://www.proteinatlas.org/>).

(less than 1% with fold change > 1.5 , $P < 0.05$; Fig. 3d) transforms to major differences between TBSCs differentiated in air-liquid interface (TBSC-ALI) with DASCs in Matrigel (DASC-MAT), consistent with their divergent fates (Fig. 3e). To further probe the differential fates of TBSC-ALI and DASC-MAT, we used laser-capture microdissection (LCM) to generate gene expression profiles of normal tracheal and alveolar epithelium and compared with those of TBSC-ALI and DASC-MAT (Fig. 3f). Gene-set enrichment analysis revealed a strong coincidence in gene expression patterns between *in vitro*-differentiated TBSCs and DASCs and their *in vivo* counterparts (FDR q value < 0.001), including many genes not previously identified as differential markers though confirmed by publically available antibody data sets²⁴ (Fig. 3g).

To determine if cloned lung stem cells could incorporate into damaged lung, we first generated single-cell-derived pedigrees of DASC^{lacZ} and TBSC^{lacZ} from murine lung (Fig. 4a; Extended data Fig. 7a). We delivered one million TBSC^{lacZ} or DASC^{lacZ} to syngeneic mice five days after influenza virus infection (Fig. 4a). At 40 dpi (35 days post-transplantation), DASC^{lacZ} were distributed in interstitial regions emanating from airways (Fig. 4b). At 90 dpi, DASC^{lacZ} showed a more homogenous pattern in interstitial spaces compared to 40 dpi (Fig. 4b). Significantly, mock-infected lungs, or mock-transplanted, infected lungs, showed

no incorporation of DASC^{lacZ} at 35 days post-transplantation (Fig. 4b; Extended data Fig. 7b). As with the lineage-tracing experiments, we used *E. coli*-specific β -galactosidase antibodies to mark the transplanted cells in 90 dpi lungs and observed that at least 40% of the cells in alveolar region express pneumocytes markers (Pdpn, 1H8 and SPC) and at least 80% of the β -galactosidase-positive cells in bronchiolar region express secretory cell marker CC10 (mouse number $n = 3$; Fig. 4c; Extended data Fig. 7c). Gene expression analysis of the lacZ-positive regions of these lungs using laser-capture microdissection revealed a typical alveoli gene signature very different from that of immature DASCs or of damaged lung (Fig. 4d). Together these findings demonstrate that single-cell-derived pedigree lines of DASCs can readily incorporate into damaged lung during the process of lung regeneration and give rise to multiple epithelial cell types of bronchioles and alveoli. In contrast, transplanted TBSC^{lacZ} appeared confined to major airways (Fig. 4e). Parallel transplantations with green fluorescent protein (GFP)-labelled DASC (Fig. 4f) yielded similar patterns of co-labelling of lineage and type I and type II pneumocyte markers at 40 dpi seen with transplanted DASC^{lacZ} (Fig. 4f). Significantly, a fraction of these transplanted DASC^{GFP} or their progeny continue to express the proliferation marker Ki67 even up to 60 dpi (Extended data Fig. 8), suggesting their high viability and extended contribution to the regenerative process. Lastly, morphometric analyses of diphtheria-toxin-treated, virally infected Krt6-DTR mice indicate that transplantation of DASCs results in a significant reduction of interstitial densities at 40 dpi (Extended data Fig. 9).

In the present work, we highlight the remarkable regenerative capacity of the lung following large-scale, acute lung damage⁷ and the function of a very discrete, pre-existing population of lung stem cells in this process. In addition, we demonstrate that upon transplantation, single-cell-derived DASC^{p63/Krt5} pedigrees contribute multiple epithelial lineages, including bronchiolar secretory cells as well as alveolar type I and type II pneumocytes, to regenerating distal lung. Thus DASC^{p63/Krt5} act in an emergent, conditional manner that is generally distinct from that of type II pneumocytes^{25–29}, progenitor cells of limited self-renewal capacity that participate in highly focal, homeostatic lung repair. Our findings provide a mechanistic framework for the still emerging concept of lung regeneration³⁰ and underscore potential therapeutic strategies exploiting this process.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 March; accepted 30 September 2014.

Published online 12 November 2014; corrected online 28 January 2015 (see full-text HTML version for details).

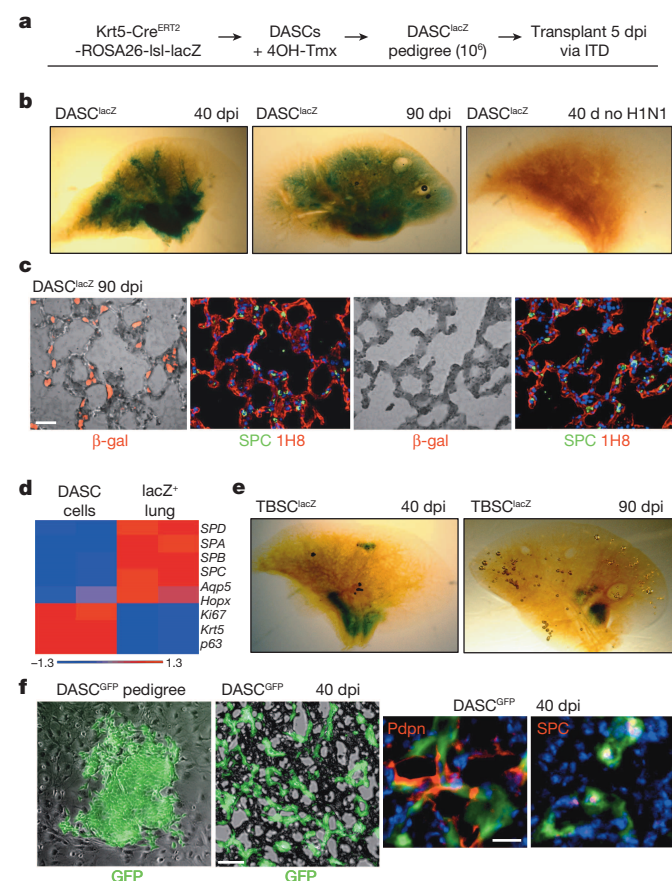


Figure 4 | Transplantation of TBSC^{lacZ} and DASC^{lacZ}. **a**, Schematic of pedigree generation and transplantation. 4OH-Tmx, 4-hydroxy-tamoxifen; ITD, intratracheal delivery. **b**, β -galactosidase activity in whole lung following DASC^{lacZ} transplantation. **c**, Comparison between β -galactosidase-positive (left panels) and -negative (right panels) regions of transplanted lung and markers of type I (1H8) and type II (SPC) cells. Scale bar, 50 μ m. **d**, Heat map of selected, differentially expressed genes ($P < 0.05$) comparing immature DASC^{lacZ} before transplantation with laser-capture microdissected lacZ-positive cells from transplanted lungs at 90 dpi. **e**, β -galactosidase activity in whole lung following TBSC^{lacZ} transplantation. **f**, From left, DASC^{GFP} colony in culture; middle, cryosection of lung following DASC^{GFP} transplantation. Scale bar, 50 μ m. Right, immunofluorescence of anti-GFP, anti-Pdpn and anti-SPC in 40 dpi transplanted lung. Scale bar, 20 μ m.

- Barnes, P. J. Chronic obstructive pulmonary disease. *N. Engl. J. Med.* **343**, 269–280 (2000).
- Gross, T. J. & Hunninghake, G. W. Idiopathic pulmonary fibrosis. *N. Engl. J. Med.* **345**, 517–525 (2001).
- Kerem, E. et al. Bacteremic necrotizing pneumococcal pneumonia in children. *Am. J. Respir. Crit. Care Med.* **149**, 242–244 (1994).
- Sawicki, G. S., Lu, F. L., Valim, C., Cleveland, R. H. & Colin, A. A. Necrotising pneumonia is an increasingly detected complication of pneumonia in children. *Eur. Respir. J.* **31**, 1285–1291 (2008).
- Herridge, M. S. et al. One-year outcomes in survivors of the acute respiratory distress syndrome. *N. Engl. J. Med.* **348**, 683–693 (2003).
- Wilcox, M. E. et al. Radiologic outcomes at 5 years after severe ARDS. *Chest* **143**, 920–926 (2013).
- Kumar, P. A. et al. Distal airway stem cells yield alveoli *in vitro* and during lung regeneration following H1N1 influenza infection. *Cell* **147**, 525–538 (2011).
- Taubenberger, J. K. & Morens, D. M. The pathology of influenza virus infections. *Annu. Rev. Pathol.* **3**, 499–522 (2008).
- Perez-Padilla, R. et al. Pneumonia and respiratory failure from swine-origin influenza A (H1N1) in Mexico. *N. Engl. J. Med.* **361**, 680–689 (2009).
- Matuschak, G. M. & Lechner, A. J. Acute lung injury and the acute respiratory distress syndrome: pathophysiology and treatment. *Mo. Med.* **107**, 252–258 (2010).
- Short, K. R., Kroeze, E. J., Fouchier, R. A. & Kuiken, T. Pathogenesis of influenza-induced acute respiratory distress syndrome. *Lancet Infect. Dis.* **14**, 57–69 (2014).
- Li, Y., Gudjonsson, J. E., Woods, T. L., Zhang, T., Johnston, A., Stoll, S. W. & Elder, J. T. Transgenic expression of S100A2 in hairless mouse skin enhances Cxcl13 mRNA

- in response to solar-simulated radiation. *Arch. Dermatol. Res.* **301**, 205–217 (2009).
13. Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nature Genet.* **21**, 70–71 (1999).
 14. Rheinwald, J. G. & Green, H. Serial cultivation of strains of human epidermal keratinocytes: the formation of keratinizing colonies from single cells. *Cell* **6**, 331–343 (1975).
 15. Barrandon, Y. & Green, H. Three clonal types of keratinocyte with different capacities for multiplication. *Proc. Natl Acad. Sci. USA* **84**, 2302–2306 (1987).
 16. Saito, M. *et al.* Diphtheria toxin receptor-mediated conditional and targeted cell ablation in transgenic mice. *Nature Biotechnol.* **19**, 746–750 (2001).
 17. Senoo, M., Pinto, F., Crum, C. P. & McKeon, F. p63 is essential for the proliferative potential of stem cells of stratified epithelia. *Cell* **129**, 523–536 (2007).
 18. Baskerville, A., Thomas, G., Wood, M. & Harris, W. J. Histology and ultrastructure of metaplasia of alveolar epithelium following infection of mice and hamsters with influenza virus. *Br. J. Exp. Pathol.* **55**, 130–137 (1974).
 19. Verhoeven, D., Teijaro, J. R. & Farber, D. L. Pulse-oximetry accurately predicts lung pathology and the immune response during influenza infection. *Virology* **390**, 151–156 (2009).
 20. Vyalov, S. L., Gabbiani, G. & Kapanci, Y. Rat alveolar myofibroblasts acquire alpha-smooth muscle actin expression during bleomycin-induced pulmonary fibrosis. *Am. J. Pathol.* **143**, 1754–1765 (1993).
 21. Tsukui, T. *et al.* Qualitative rather than quantitative changes are hallmarks of fibroblasts in bleomycin-induced pulmonary fibrosis. *Am. J. Pathol.* **183**, 758–773 (2013).
 22. Lawson, W. E. *et al.* Characterization of fibroblast-specific protein 1 in pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **171**, 899–907 (2005).
 23. Rock, J. R., Randell, S. H. & Hogan, B. L. Airway basal stem cells: a perspective on their roles in epithelial homeostasis and remodeling. *Dis. Model. Mech.* **3**, 545–556 (2010).
 24. Schmidt, D., Hubsch, U., Wurzer, H., Heppt, W. & Aufderheide, M. Development of an *in vitro* human nasal epithelial (HNE) cell model. *Toxicol. Lett.* **88**, 75–79 (1996).
 25. Uhlen, M. *et al.* Towards a knowledge-based Human Protein Atlas. *Nature Biotechnol.* **28**, 1248–1250 (2010).
 26. Evans, M. J., Cabral, L. J., Stephens, R. J. & Freeman, G. Transformation of alveolar type 2 cells to type 1 cells following exposure to NO₂. *Exp. Mol. Pathol.* **22**, 142–150 (1975).
 27. Sugihara, H., Toda, S., Miyabara, S., Fujiyama, C. & Yonemitsu, N. Reconstruction of alveolus-like structure from alveolar type II epithelial cells in three-dimensional collagen gel matrix culture. *Am. J. Pathol.* **142**, 783–792 (1993).
 28. Barkauskas, C. E. *et al.* Type 2 alveolar cells are stem cells in adult lung. *J. Clin. Invest.* **123**, 3025–3036 (2013).
 29. Desai, T. J., Brownfield, D. G. & Krasnow, M. A. Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* **507**, 190–194 (2014).
 30. Xian, W. & McKeon, F. Adult stem cells underlying lung regeneration. *Cell Cycle* **11**, 887–894 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by grants from the Joint Council Office of the Agency for Science Technology Research Agency (ASTAR), Singapore (W.X., F.M.), Defense Advanced Research Projects Agency (DARPA, N66001-09-1-2121 to F.M.), the Johnson & Johnson ASTAR Joint Program (W.X., F.M.) and support from Connecticut Innovations (W.X., F.M.). We thank M. LaLande, B. Lane and H. Hui Ng for support, G. Wright, B. Tennent, B. Knowles and T. McLaughlin for comments on the manuscript, J. Hammer for artwork, P. Kraus for blastocyst injections, and H. Ahmad and K. L. Goh for technical assistance. We thank H. Green for advice and support.

Author Contributions Experiments were performed by W.Z., D.Z.W., S.P.G. and A.-A.L. Experimental design and conception were done by W.Z., M.V., C.P.C., W.X. and F.M.; T.Z., W.Z., X.W., S.J.L. and Y.Y. performed microarrays and computational analysis, and provided methodological advice. M.L. and W.Z. performed the serial reconstructions of infected lung.

Author Information Datasets generated for this study have been submitted to the National Center for Biotechnology Information Gene Expression Omnibus (GEO) database under superseries GSE60849. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.M. (mckeon.xian@gmail.com) or W.A. (waxian@jax.org).

METHODS

Influenza virus infection. All mouse experiments were conducted under IACUC guidelines and approved protocols. Influenza A (H1N1) mouse-adapted PR/8/34 (VR-95, ATCC, USA) was used for all viral infections. The virus stock was amplified by V. Chow (Department of Microbiology, National University of Singapore) in chicken eggs. Virus dilutions were made in DMEM medium containing $1 \mu\text{g ml}^{-1}$ TPCK trypsin (Sigma-Aldrich, USA) on ice, aliquoted and stored at -80°C . The viral titre is measured by plaque assay on Madin Darby canine kidney cells (MDCK, ATCC, USA). Virus was further diluted to final concentration in PBS on ice and use freshly. The infection of mice by H1N1 Influenza virus was performed in an Animal Biosafety Level 2 (ABSL-2) facility. Adult mice (>6 weeks old) were anaesthetized with intraperitoneal injection of ketamine (150 mg per kg body weight) + xylazine (10 mg per kg body weight). The anaesthetized mouse was rested on a stand with its front teeth hung over a suture. This causes the mouse airway to be relaxed and accessible. Using a flat forceps the tongue of the animal was drawn out of its mouth so that the anatomy can be easily visualized. Intratracheal delivery was performed by pipetting $50 \mu\text{l}$ virus directly into the larynx/trachea. Sterile PBS was administered to the control animals. The tongue of mouse was held throughout the procedure so that the virus was aspirated into the lungs.

Tissue histology. At appropriate time points, mice were euthanized by CO_2 asphyxiation followed by exsanguination, and the diaphragm was carefully cut open without touching the lungs. A small incision was made in the proximal region of the trachea and lung was inflated with 4% formaldehyde using a 30G needle. The inflated lungs were dissected and fixed with 4% formaldehyde before whole mount imaging, paraffin section or cryosection. For whole mount imaging, lungs were dehydrated in graded ethanol series and sunk in BABB (benzyl alcohol/benzyl benzoate 1:2 ratio) at 4°C overnight³¹. For paraffin section, lungs were processed in an automatic tissue processor (Leica Microsystems, Germany) and embedded into paraffin blocks. The blocks were cut using microtome (Leica Microsystems, Germany) to $5\sim 7 \mu\text{m}$ thickness at distinct planes. The sections were placed on poly-lysine coated glass slides and stored at room temperature until further use. For cryosection, lungs were embedded within Tissue-Tek O.C.T compound, solidified on dry ice and cut using a cryotome (Leica Microsystems, Germany) of $10 \mu\text{m}$ thickness.

Haematoxylin and eosin (H&E) staining was performed using standard procedures. To analyse lung damage level, interstitial densities were assessed by H&E staining backed up by type I (anti-Pdpn) and type II (anti-SPC) pneumocytes staining. A minimum of 8 axial lung interval sections (typically 400 mm^2) covering $>2 \text{ mm}$ tissue depth were cut, each was stained, scored for densities, and quantified for percentage of total lung area by Zeiss AxioVision (Carl Zeiss, Germany) morphometric software. In addition, random histological sections are scored based on the general pathological morphology by blinded expert to confirm the conclusion. Stitching scanning of H&E slides were performed in histopathology lab of IMCB, A-STAR, Singapore. Masson Trichrome staining of lung fibrosis was performed using the Trichrome Staining Kit (Sigma-Aldrich, USA). The kit involves sequential staining of the sections with Biebrich Scarlet-Acid Fuchsin, PTA/PMA and Aniline Blue. After staining, sections were dehydrated and mounted using Vectamount and visualized under a light microscope (Imager Z1, Carl Zeiss, Germany).

Immunofluorescence staining. For immunofluorescence staining, paraffin-embedded tissue slides were subjected to antigen retrieval in citrate buffer (pH 6, Sigma-Aldrich, USA) at 120°C for 20 min with the exception of CD45 staining. Antibodies used for immunofluorescence included stem cell markers: Krt5 (1:200, EP1601Y, Thermo), Krt6 (1:100, T-18, Santa Cruz and Ab24646, Abcam), p63 (1:2, 4A4 clone, house-made); pneumocyte markers: Pdpn (1:100, M-172 and A-18, Santa Cruz), Aqp5 (1:100, G-19 and H-200, Santa Cruz), Hoxp (1:100, FL-73, Santa Cruz), SPC (1:100, M-20 and FL-197, Santa Cruz), 1H8 mouse monoclonal antibodies (1:2, house-made); others: CC10 (1:100, T-18, Santa Cruz), CD45 (1:100, 30-F11, Santa Cruz), bacteria-specific β -galactosidase (1:400, A-11132, Life Technologies and Ab9361, Abcam), Ki67 (1:200, RM-9106, Thermo), α -SMA (1:400, 1A4, Dako), HB-EGF (1:100, Ab16783, Abcam), GFP (1:100, B-2, Santa Cruz) and acetylated α tubulin (1:1,000, Ab24610, Abcam). Among them, 1H8 murine monoclonal antibody was generated under IRB approval using standard methods, and validated for specific mouse type I pneumocyte staining by immunofluorescence. Alexa-conjugated secondary antibodies (1:200) were used for immunofluorescence. After staining, tissues slides underwent auto-fluorescence removal and mounting with DAPI containing mounting media (Vectashield, Vector Labs, USA). Stained slides were stored at 4°C in the dark and images were taken using Zeiss fluorescence microscope (Observer Z1, Carl Zeiss, Germany) or Zeiss confocal microscope (LSM 510, Carl Zeiss, Germany).

Lineage tracing of Krt5⁺ stem cells. Krt5-CRE/ERT2 (011916-MU, MMRRRC, USA) mice were crossed with Rosa26-loxP-STOP-loxP-lacZ mice (003309, Jackson Laboratory, USA) to generate mice for lineage tracing. Genotype was confirmed

for each mouse using tail genomic DNA collection and PCR validation. Tamoxifen was dissolved in corn oil and freshly applied to mice at 200 mg per kg body weight through intraperitoneal injection at indicated days before influenza infection. The gap between tamoxifen and H1N1 administration was varied as indicated to control for any possible tamoxifen persistence. Sublethal doses (25 plaque forming units) of H1N1 influenza A virus were diluted in PBS and intratracheally delivered into anaesthetized mice. After infection by virus, mouse lungs were collected at various time points and subjected to standard immunofluorescence staining by two independent bacteria-specific β -galactosidase antibodies (1:400, A-11132, Life Technologies and Ab9361, Abcam) or X-gal staining. For X-gal staining, lungs were briefly fixed on ice for 30 min and subjected to X-gal (Invitrogen, USA) whole-mount staining overnight using standard protocol. After staining, lungs were washed and fixed again in 4% formaldehyde before whole-mount visualization or paraffin sectioning. For whole-mount visualization, lungs were made transparent by BABB as described before and images were taken using dissection microscope (Leica Microsystems, Germany). For paraffin section, $5\sim 7 \mu\text{m}$ sections were cut and immunofluorescence staining is performed following standard protocol and visualized under light microscope (Imager Z1, Carl Zeiss, Germany).

Generation of the Krt6-DTR mouse. The complementary DNA of diphtheria toxin receptor (DTR), which is also known as human heparin-binding epidermal growth factor-like growth factor (HB-EGF), and a neomycin resistance selection cassette flanked with loxPs (Floxed Neo^R) with an introduced PacI restriction endonuclease site, were introduced to replace the first 4 exons of Krt6a in a modified bacterial artificial chromosome (mBAC). Retrieval and linearization of a selected section of this mBAC resulted in the targeting construct, which was electroporated into V6.4 B6.129 hybrid embryonic stem cells to be selected with G418. Single colonies screened by Southern blot analysis of PacI digests detected by a 5' external probe (hybridizing unmodified Krt6a genomic DNA) revealed fragment size differences due to the introduced PacI site. Wild-type alleles with endogenous PacI sites returned 35.9 kb fragments, while a recombination event returned a shorter 14.7 kb. Floxed Neo^R probes provided further verification for a single specific insertion. Successfully engineered embryonic stem cells were micro-injected into blastocysts to generate chimaeras, which were similarly tested for germline transmission via backcrosses to C57BL/6. Progeny from crosses with FVB/N-Tg(ACTB-cre)2Mrt/J (Jackson Laboratory, USA) were screened for stable transmission of Krt6a-DTR alleles. Cre-mediated excision of Floxed Neo^R as indicated by a reduced 12.9 kb PacI digest fragment.

Oxygen saturation measurements. Peripheral capillary oxygen saturation (SpO₂) was measured using MouseOx Plus pulse oximeter (Starr Life Sciences, USA). An S-size CollarClip sensor was applied to depilated regions of the back neck skin and mice were rested for one hour before measuring SpO₂. Ten minutes before measurement, mice were anaesthetized by ketamine (150 mg per kg body weight) + xylazine (10 mg per kg body weight) intraperitoneal injection. After SpO₂ readings were stable, data collection was started and SpO₂ readings were recorded every second for one minute to calculate an average value.

Cloning of TBSC and DASC and *in vitro* differentiation. To isolate airway stem cells, trachea and lung were collected from adult mice and immersed in cold wash buffer (F12 medium, 1% Pen/Strep, 5% FBS). The trachea and two main bronchi were separated from the lungs and the lobes were cut with a sterile surgical blade into small pieces and digested with dissociation buffer (F12/DMEM, 1 mg ml^{-1} protease, 0.005% trypsin and 10 ng ml^{-1} DNase I) overnight with gentle rocking. The dissociated cells were washed with wash buffer, passed through a $40 \mu\text{m}$ cell strainer, counted and plated onto irradiated 3T3 feeder cells as described^{7,14,15}. After 4 consecutive passages, single cell colonies were picked up by cloning ring and expanded. Colonies were characterized by immunofluorescence staining (E-cadherin⁺ Krt5⁺ p63⁺ Pdpn⁺ CC10⁺ SPC⁺). These colonies have been passaged up to 12 months with no observable phenotypic or chromosome count changes. To compare gene expression in immature TBSCs and DASCs, whole genome expression microarrays were performed on passage 7 (P7) cells. Figure 3c is a scatterplot comparison of the TBSC and DASC whole transcriptome data. Gene names are indicted for outliers (fold change > 3 , P value < 0.001) typically reported as markers or well known for other biology and for which we validated expression by quantitative PCR. To activate lacZ in TBSCs and DASCs cloned from Rosa26-lsl-lacZ; Krt5-Cre^{ERT2} mice, we exposed colonies *in vitro* to 4-OH-tamoxifen ($1 \mu\text{g ml}^{-1}$; Sigma-Aldrich) for 5 days at which point $> 70\%$ cells express β -galactosidase and all cells were Krt5⁺ p63⁺ SPC⁺ CC10⁺ Pdpn⁺. GFP labelling was performed by retroviral transduction of pMX-GFP (Cell Biolab, USA) followed by manual sorting of single GFP⁺ DASCs to 96-well plates.

Matrigel differentiation assays were performed as previously described⁷. FGF10 (50 ng ml^{-1}) was included in medium to favour distal airway differentiation. Under this condition, DASCs clustered and grew into sphere-like structures marked by unilaminar epithelia surrounding a clear lumen. TBSCs were grown on air-liquid interface (ALI) cultures for differentiation⁷. ALI differentiation was performed as

previously described¹¹, FGF10 was excluded and retinoid acid (50 nM) included in medium to favour proximal airway differentiation. Under ALI condition, TBSC forms a stratified epithelium structure marked by ciliated cells (acetylated-Tub+) and basal cells (Krt5⁺). In addition, we used laser capture microdissection (LCM, Zeiss PALM) to dissect mouse trachea epithelium, alveoli and damaged lung interstitium for microarray analysis to develop tissue specific gene expression signatures.

Orthotopic transplantation of stem cells. To perform orthotopic transplantation of stem cells into lung, adult mice were infected with 25 plaque forming units H1N1 influenza virus five days prior transplantation. Stem cell pedigrees were expanded in culture and harvested by differential trypsinization to remove feeder cells. One million cells were diluted in 50 µl DMEM/F12 medium for transplantation into each mouse. Adult mice (>6 weeks old) were anaesthetized with intraperitoneal injection of ketamine (150 mg per kg body weight) + xylazine (10 mg per kg body weight) and rested on a stand gesture. Intratracheal aspiration was performed by pipetting the virus directly into trachea via mouth. For rescue of Krt6-DTR phenotype by DASC transplantation, 3×10^5 cells were transplanted on 8 dpi (the same day DTox was given) and control mice received medium alone.

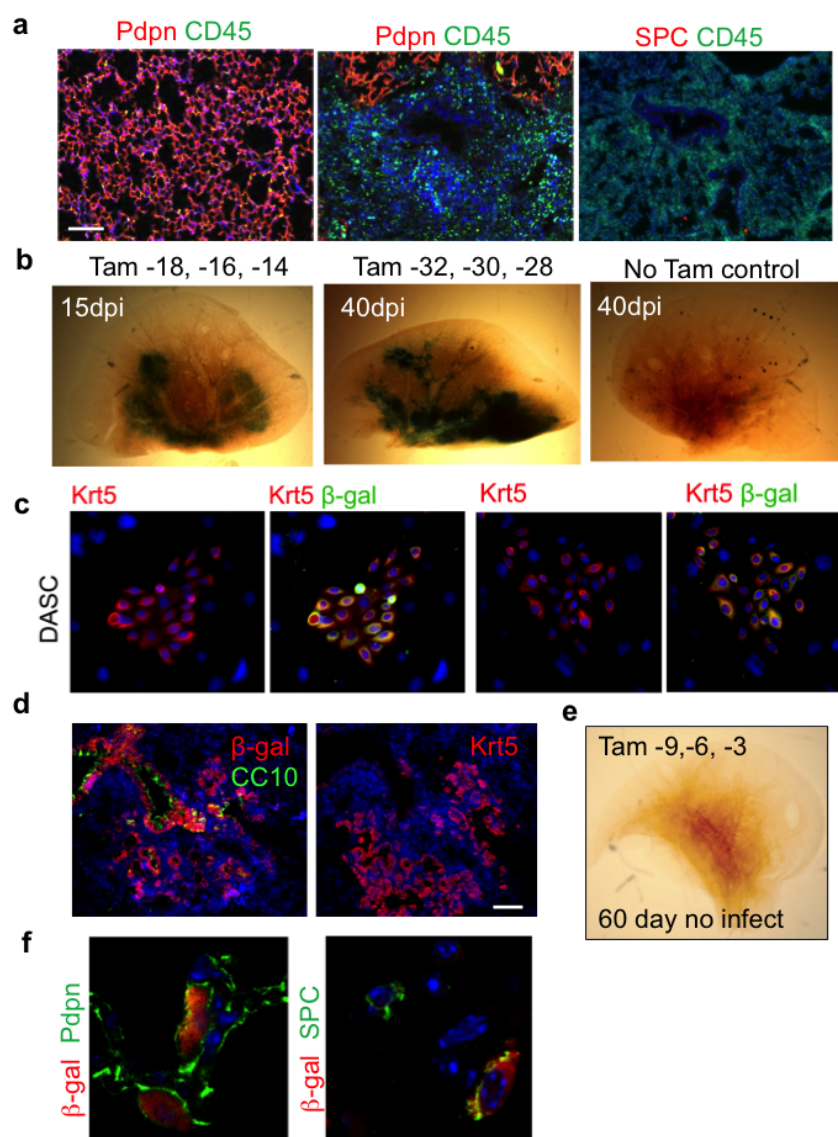
Laser capture microdissection (LCM). For LCM, fresh, non-fixed or X-gal stained tissue samples were embedded in Tissue-Tek O.C.T. compound (Sakura, Japan) on dry ice. 10 µm cryosections were mounted on PEN Membrane slides (Leica Microsystems, Germany). Slides were dehydrated in 95%, 75%, 50% nuclease-free ethanol for 30 s each, stained in Arcturus HistoGene staining solution (Life Technologies, USA) for 1 min then dehydrated in 50%, 75%, 95%, 100% ethanol for 30 s each and xylene for 5 min. Slides were allowed to air dry and LCM was performed immediately using an inverted microscope and PALM Robo software (Carl Zeiss, Germany). Cut elements were catapulted onto an Adhesive Cap-500 tube (Carl Zeiss) and subsequently transferred into a PCR tube with 50 µl Arcturus PicoPure extraction buffer. RNA was isolated using the Arcturus PicoPure RNA Isolation kit (Life Technologies, USA).

Microarray and bioinformatics. RNAs obtained from LCM, cell colonies or whole lobe of mouse lungs were used for microarray after being amplified using WT-Ovation Pico RNA Amplification System (NuGEN, UK) and fragmented and labelled using the FL-Ovation cDNA Biotin Module V2 (NuGEN, UK). Labelled cDNA was then hybridized onto GeneChip Mouse Exon 1.0 ST Array (Affymetrix, USA) using appropriate hybridization controls and the chips were scanned and

analysed as described previously⁷. Duplicate experiments for microarray were taken from two biological samples. To validate sample quality, probe hybridization ratios were calculated using Affymetrix Expression Console software (Affymetrix, USA). The intensity values were log₂-transformed and imported into the Partek Genomics Suite 6.6 (Partek Inc., USA). Exons were summarized to genes and 1-way ANOVA was performed to identify differentially expressed genes. P values and fold-change were calculated for each analysis. Heat maps were generated using Pearson's correlation and Ward's method and principal component analysis was conducted using all probe sets. Gene ontology analyses were performed using the web-based GeneTrail tool³².

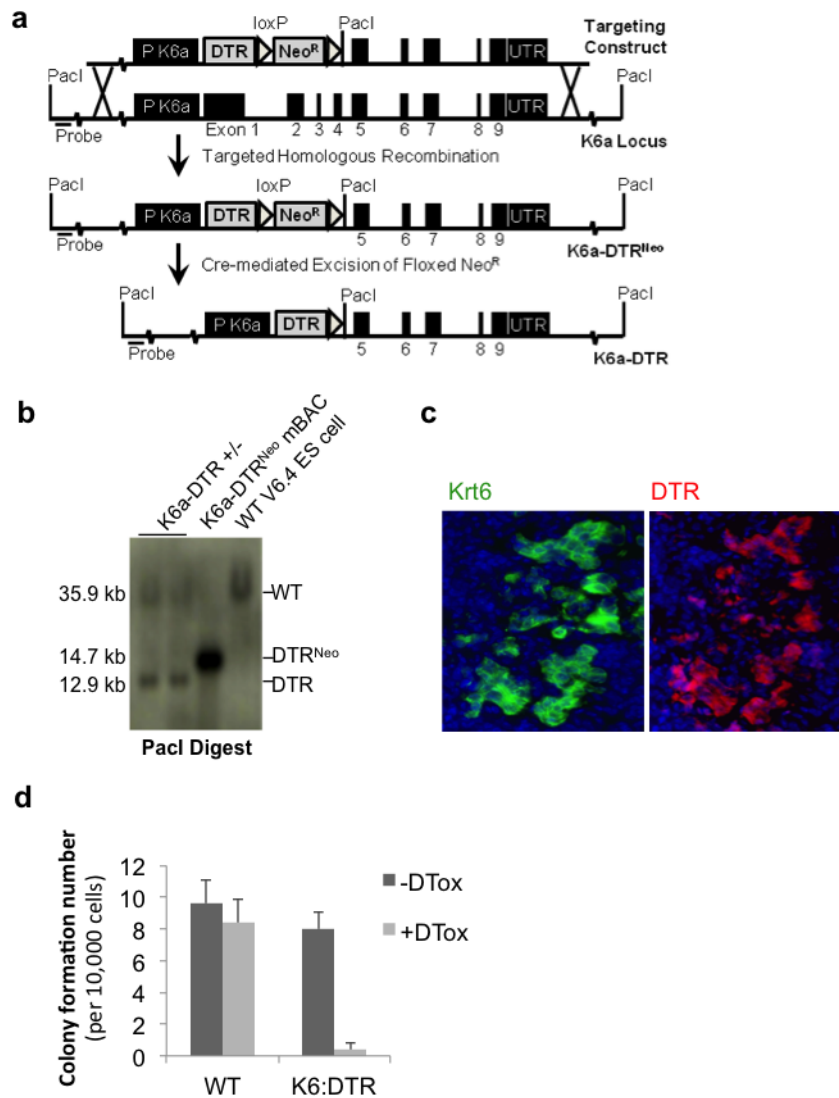
For the comparison between gene expression profiles of *in situ* alveoli and tracheal epithelium with those of *in vitro* differentiated TBSCs and DASCs (see Fig. 3f, g), gene set enrichment analyses were performed using web-based tools developed by the Broad Institute^{33,34}. 605 genes highly expressed in tracheal epithelium (trachea versus alveoli, 4-fold higher, *P* value < 0.05) and 914 genes highly expressed in alveolar epithelium (trachea versus alveoli, 4-fold lower, *P* value < 0.05) were used as the queried gene sets (signature), respectively. Enrichment scores were determined after 1,000 permutations, and the permutation type was configured to the gene sets. Thereafter the whole genome expression data of DASC-MAT versus TBSC-ALI were applied to GSEA program to evaluate enrichment of signature in the fold-change ordered list. Results with normalized enrichment score > 1.4 and FDR *q* value < 0.001 were considered significant. The comparison between gene expression profiles of normal alveoli and damaged interstitium with those of lung tissue with or without DASC ablation (see Extended Data Fig. 3a) was performed in similar manner.

31. Becker, K., Jährling, N., Saghafi, S. & Dodt, H. U. Immunostaining, dehydration, and clearing of mouse embryos for ultramicroscopy. *Cold Spring Harb. Protoc.* **2013**, 743–744 (2013).
32. Keller, A. *et al.* GeneTrailExpress: a web-based pipeline for the statistical evaluation of microarray experiments. *BMC Bioinformatics* **9**, 552 (2008).
33. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
34. Mootha, V. K. *et al.* PGC-1 alpha responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genet.* **34**, 267–273 (2003).



Extended Data Figure 1 | Lineage tracing of Krt5⁺ cells. **a**, Left, immunofluorescence images of sections of 15 dpi lung with staining patterns of antibodies to pan-leukocyte marker CD45 and the type I pneumocyte marker Pdpn with DNA counterstained with DAPI. Right, immunofluorescence images of pan-leukocyte marker CD45 and the type II pneumocyte marker SPC. Scale bar, 150 μ m. **b**, X-gal staining (blue) to reveal lacZ-dependent β -galactosidase activity in whole lungs after 15 and 40 days post infection following long time gaps between induction of lacZ labelling by tamoxifen and influenza infection-induced lung damage. The similarity of this long gap labelling and the short gap labelling presented in Fig. 1 argues against prolonged actions of tamoxifen in these lineage-labelling protocols. Tamoxifen is given at

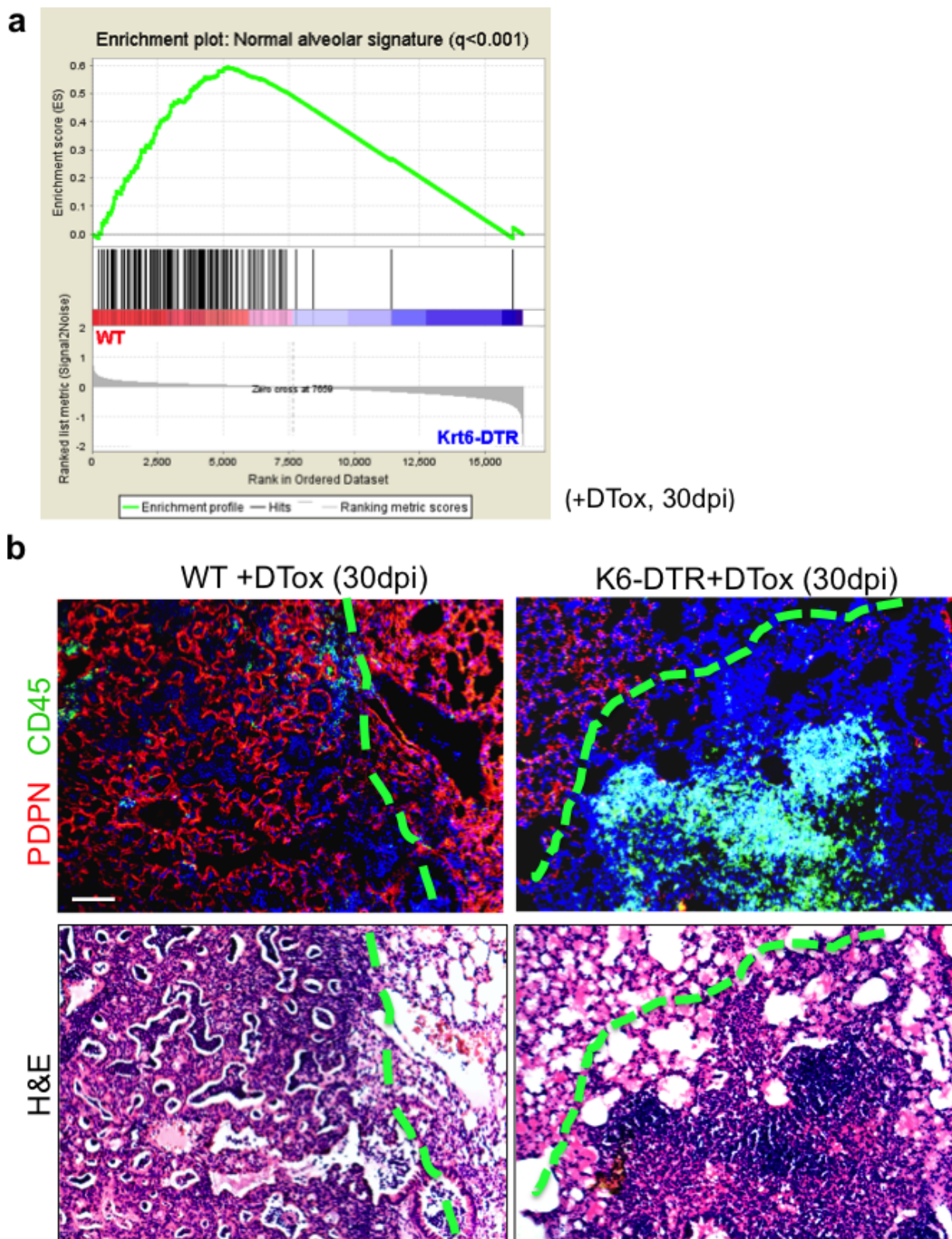
indicated times before infection and no-tamoxifen control is included. **c**, Immunofluorescence images of colonies of DASCs derived from tamoxifen-treated ROSA26-lsl-lacZ; Krt5-Cre^{ERT2} mice stained with antibodies to keratin 5 (Krt5) or Krt5 and *E. coli*-specific β -galactosidase. **d**, Histological section of lung at 15 dpi stained with *E. coli*-specific β -galactosidase antibody and markers of secretory cells (CC10⁺) and expanded stem cells (Krt5⁺). Scale bar, 50 μ m. **e**, Whole-mount image of X-gal developed, uninfected lung from ROSA26-lsl-lacZ; Krt5-Cre^{ERT2} which received tamoxifen treatments at -69, -66 and -63 days before dissection. **f**, Histological section of 60 dpi lung stained with *E. coli*-specific β -galactosidase antibody and markers of type I pneumocytes (Pdpn⁺) and type II pneumocytes (SPC⁺).



Extended Data Figure 2 | Conditional DASC^{p63/Krt5} ablation mouse model.

a, Schematic of *Krt6a* locus, the targeting vector constructed to introduce the human diphtheria toxin receptor (DTR). **b**, The structure of the modified *Krt6a* locus in embryonic stem cells screened by Southern blot.

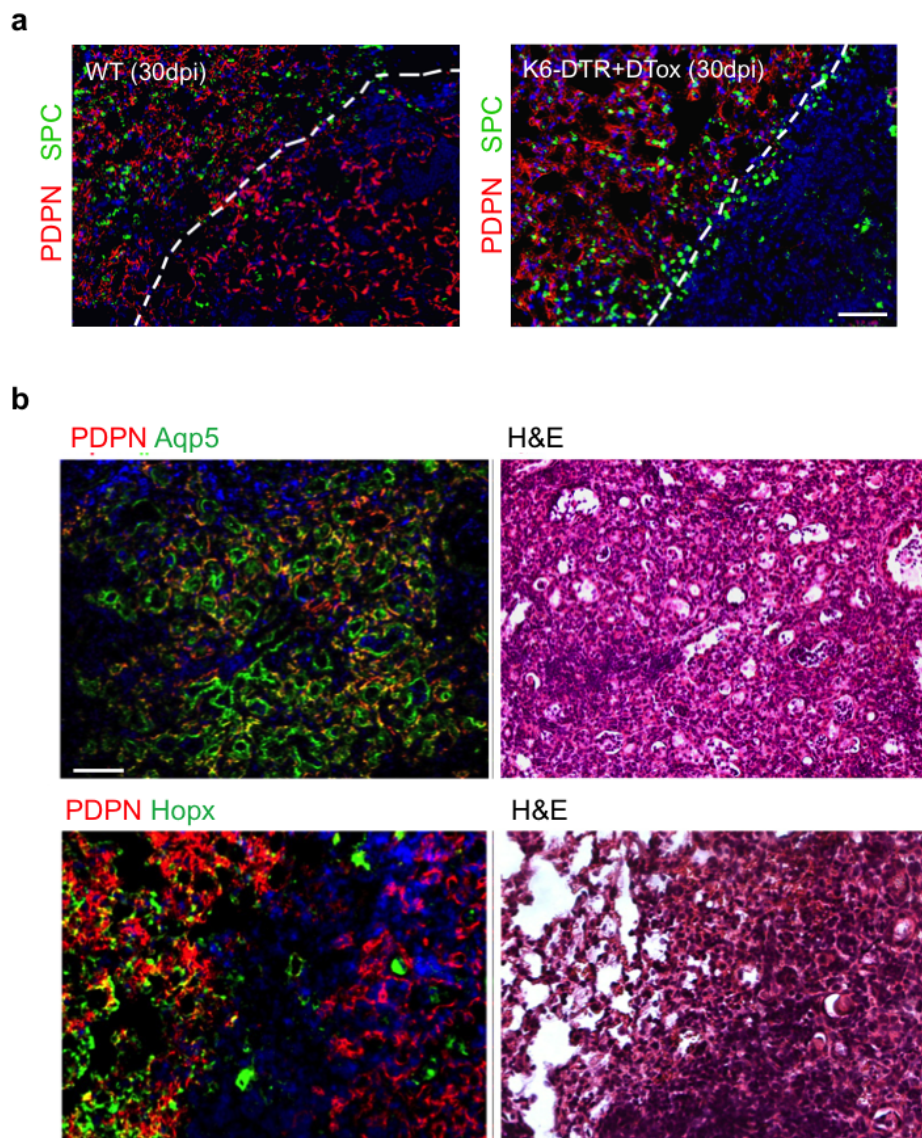
c, Co-expression of Krt6 and DTR in Krt5⁺ pods in 15 dpi lung. **d**, Histogram showing resistance of wild-type, 12 dpi DASC^{p63/Krt5} to diphtheria toxin (DTox) and the sensitivity of DASC^{p63/Krt5/DTR} to diphtheria toxin. $n = 3$ mice per group. Error bars, s.e.m.



Extended Data Figure 3 | Persistent damage in DASC^{p63/Krt5}-ablated lungs.

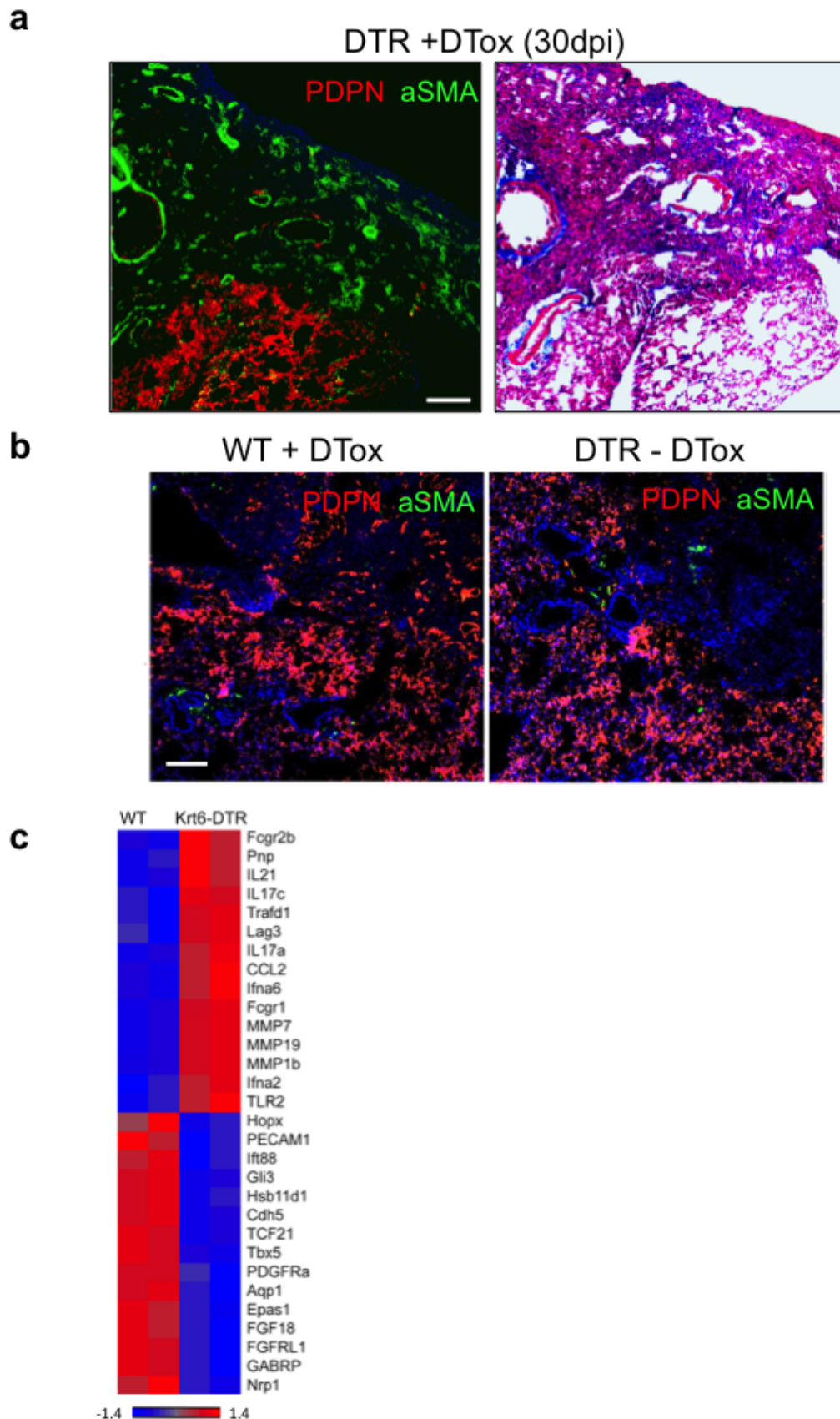
a, Gene set enrichment analysis (GSEA) showing the overrepresentation of normal alveolar signature gene sets in WT rather than Krt6-DTR mouse lungs (whole lobes). For normal alveolar signature build up, laser capture microdissection of frozen sections was used to dissect normal alveoli region from 0 dpi lung and damaged interstitial infiltrated region from 15 dpi lung for microarray analysis. Differentially expressed genes (fold change > 5 , $P < 0.01$) were used to develop normal alveolar gene expression signatures. **b**, Top

panel, histological analysis of lung densities using anti-Pdpn antibodies (red) and anti-CD45 (green) to reveal type I pneumocytes and leukocyte infiltration, respectively. Left, wild-type mice showing apparently normal lung region adjacent to interstitial density having Pdpn⁺ network but lacking CD45⁺ infiltrates. Right, Krt6-DTR lung showing apparently normal region adjacent to zone of damaged interstitial lung lacking Pdpn⁺ network but having CD45⁺ infiltrates. Bottom panel, H&E staining of the same histological region. Scale bar, 100 μ m.



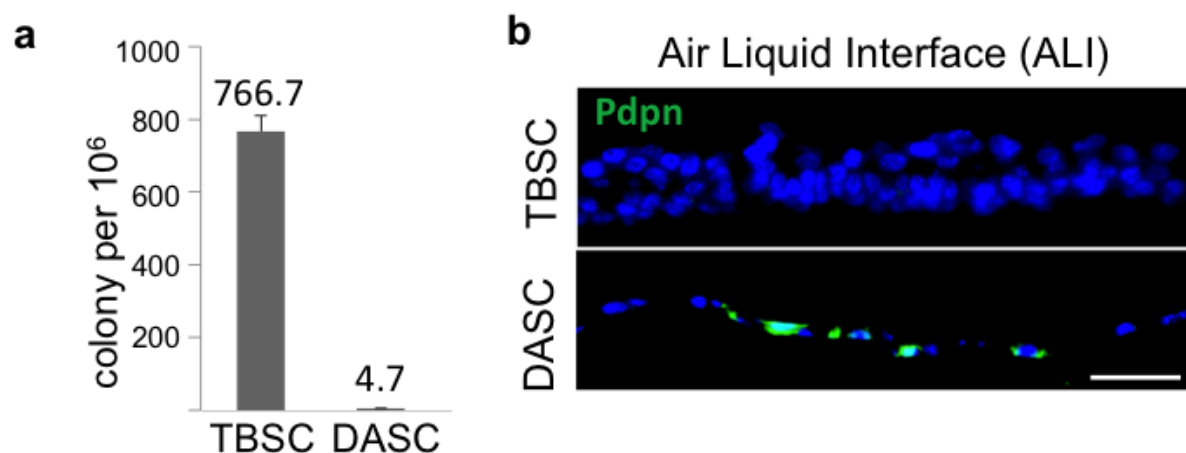
Extended Data Figure 4 | Networks of type I pneumocytes in 30 dpi mouse lung. **a**, Histological analysis of lung densities using anti-Pdpn antibodies (red) and anti-SPC (green) to reveal type I and type II pneumocytes respectively. Left, wild-type mice showing apparently normal lung region adjacent to interstitial density having Pdpn⁺ network but lacking SPC⁺ cells. Right, Krt6-DTR lung showing normal region adjacent to zone of damaged interstitial lung lacking both pneumocytes. Scale bar, 100 μ m. **b**, Top panel, histological

analysis of wild-type lung densities using anti-Pdpn antibodies (red) and anti-Aqp5 (green) type I pneumocyte markers showing the interstitial density having Pdpn/Aqp5 double-positive network. Bottom panel, wild-type mice show apparently normal lung region adjacent to interstitial density having Pdpn⁺ network but the density lack expression of another type I pneumocyte marker, Hopx. Scale bar, 100 μ m.



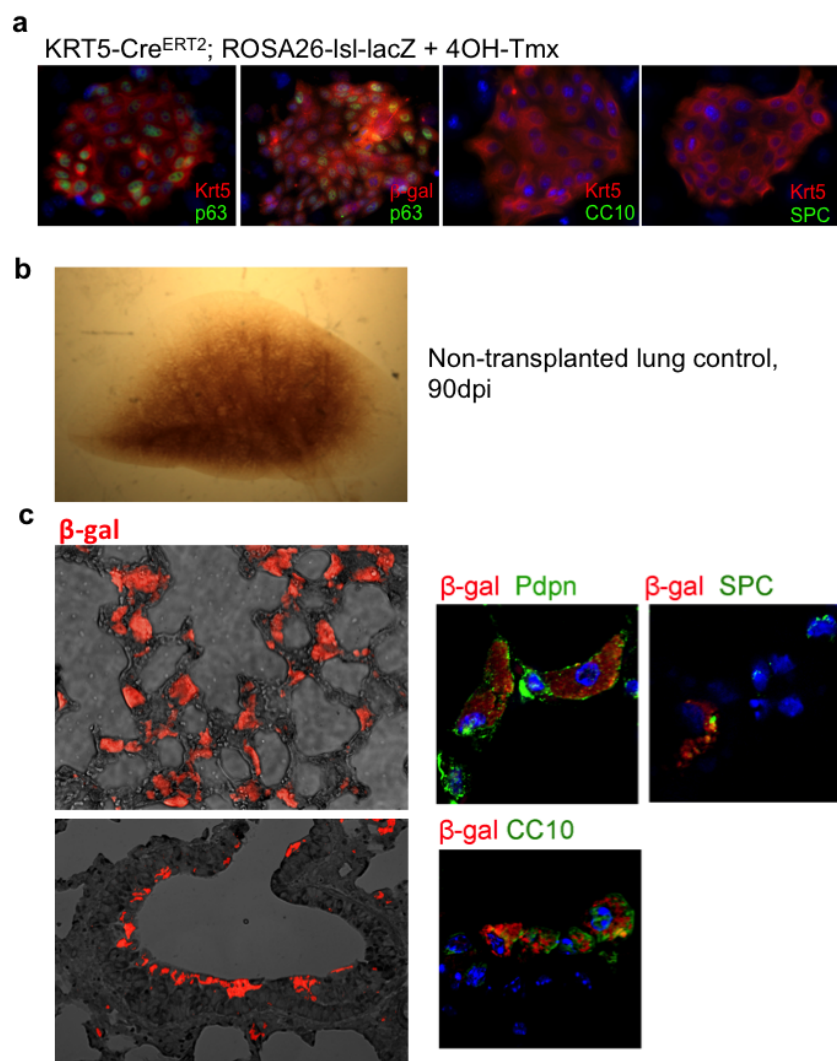
Extended Data Figure 5 | Failure of regeneration in DASC^{p63/Krt5}-ablated lungs. **a**, Histological section through 30 dpi DASC-ablated lung (Krt6-DTR +DTox) showing normal region (Pdpn⁺) adjacent to interstitial density positive for α-SMA and weakly positive for Masson's trichrome (MT) staining for fibrosis. Scale bar, 100 μm. **b**, Histological section through 30 dpi control

lung showing normal region (Pdpn⁺) and interstitial density (Pdpn⁻) which are both negative for α-SMA. **c**, Expression heat map of selected, differentially expressed genes ($P < 0.05$) comparing wild-type mouse lungs with DASC-ablated mouse lungs at 30 dpi. Scale bar, 100 μm.



Extended Data Figure 6 | Cloning and *in vitro* differentiation of TBSC^{p63/Krt5} and DASC^{p63/Krt5}. **a**, Histogram of cloning efficiency of TBSCs and DASCs on irradiated 3T3-J2 cells per 1 million tracheal or distal airway cells derived from respective tissues of adult mice. Tissues derived from 3 mice.

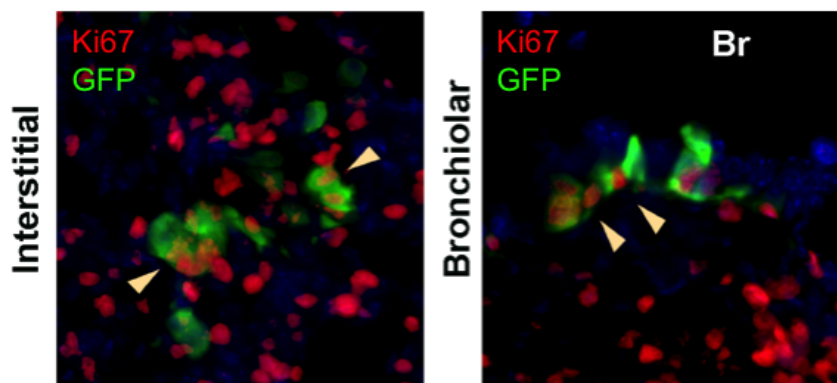
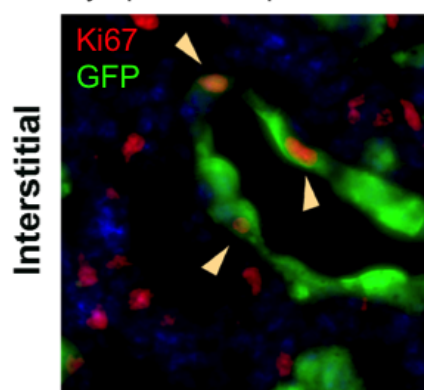
Error bars, s.e.m. **b**, Immunofluorescence images of sections of TBSC and DASC air–liquid interface cultures using an antibody to the type I pneumocyte marker Pdpn (green). Sections were counterstained with DAPI (blue).



Extended Data Figure 7 | Transplantation of DASC^{lacZ}.

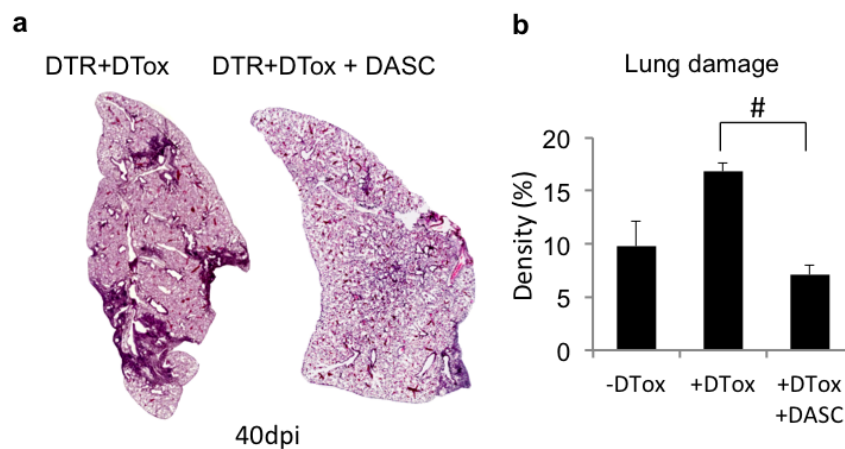
a, Immunofluorescence characterization of DASCs isolated from Krt5-Cre^{ERT2}; ROSA26-lsl-lacZ mice following Cre activation with 4OH-tamoxifen. From left, colony stained with antibodies to p63 (green) and Krt5 (red), p63 (green) and *E. coli* β-galactosidase (red), Krt5 (red) and CC10 (green), and Krt5 (red) and SPC (green). **b**, Whole mount image of lung 90 days after

infection without stem cell transplantation. **c**, Left, bright field/immunofluorescence image of section of lung at 90 dpi following transplantation of DASC^{lacZ} stained with antibodies to β-galactosidase (red). Right panels, immunofluorescence images of co-staining of transplanted DASC^{lacZ} with antibodies to Pdpn, SPC, or CC10 at high magnification.

7 days post-transplantation DASC^{GFP}55 days post-transplantation DASC^{GFP}

Extended Data Figure 8 | Persistent proliferation of transplanted DASC. Co-staining of antibodies to GFP (green) with the cell proliferation marker Ki67 (red) in sections of lung transplanted with DASC^{GFP} at 12 dpi lung (7 days post transplantation) and 60 dpi lung (55 days post transplantation). Top left, immunofluorescence image of lung following transplantation of DASC^{GFP} (7 days post-transplantation; 12 dpi) stained with anti-GFP (green)

and the cell cycle marker Ki67 (red, in nucleus). Top right, bronchiole co-stained with antibodies to GFP and Ki67 from 7 days post-transplantation lung. Bottom, staining of interstitial lung transplanted 55 days prior with DASC^{GFP} with antibodies to GFP and Ki67. Arrows indicate cells co-expressing GFP and Ki67.



Extended Data Figure 9 | Stem cell transplantation reduces interstitial densities in DASC^{p63/Krt5}-ablated lungs. **a**, Histological sections through entire lobe of Krt6–DTR mice with (left) and without (right) diphtheria toxin treatment forty days post-influenza infection. **b**, Histogram of morphometric quantification of lung densities following 40 day influenza virus infection of

Krt6–DTR mice without diphtheria toxin (–DTox, mouse number $n = 3$), with diphtheria toxin (+DTox, $n = 4$), or with diphtheria toxin and transplanted DASCs (+DTox+DASC, $n = 4$). Error bars indicate s.e.m. and # indicates P value = 0.029 by Wilcoxon rank-sum test.

Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury

Andrew E. Vaughan¹, Alexis N. Brumwell¹, Ying Xi¹, Jeffrey E. Gotts¹, Doug G. Brownfield², Barbara Treutlein³, Kevin Tan¹, Victor Tan¹, Feng Chun Liu¹, Mark R. Looney¹, Michael A. Matthay¹, Jason R. Rock⁴ & Harold A. Chapman¹

Broadly, tissue regeneration is achieved in two ways: by proliferation of common differentiated cells and/or by deployment of specialized stem/progenitor cells. Which of these pathways applies is both organ- and injury-specific^{1–4}. Current models in the lung posit that epithelial repair can be attributed to cells expressing mature lineage markers^{5–8}. By contrast, here we define the regenerative role of previously uncharacterized, rare lineage-negative epithelial stem/progenitor (LNEP) cells present within normal distal lung. Quiescent LNEPs activate a Δ Np63 (a p63 splice variant) and cytokeratin 5 remodelling program after influenza or bleomycin injury in mice. Activated cells proliferate and migrate widely to occupy heavily injured areas depleted of mature lineages, at which point they differentiate towards mature epithelium. Lineage tracing revealed scant contribution of pre-existing mature epithelial cells in such repair, whereas orthotopic transplantation of LNEPs, isolated by a definitive surface profile identified through single-cell sequencing, directly demonstrated the proliferative capacity and multipotency of this population. LNEPs require Notch signalling to activate the Δ Np63 and cytokeratin 5 program, and subsequent Notch blockade promotes an alveolar cell fate. Persistent Notch signalling after injury led to parenchymal ‘micro-honeycombing’ (alveolar cysts), indicative of failed regeneration. Lungs from patients with fibrosis show analogous honeycomb cysts with evidence of hyperactive Notch signalling. Our findings indicate that distinct stem/progenitor cell pools repopulate injured tissue depending on the extent of the injury, and the outcomes of regeneration or fibrosis may depend in part on the dynamics of LNEP Notch signalling.

Influenza infection challenges pulmonary regenerative capacity owing to the widespread ablation of epithelial cells in substantial areas of lung⁸ (Extended Data Fig. 1g, h). A robust expansion of regenerative cytokeratin-5-positive (Krt5⁺) cells in the lung parenchyma after influenza infection has been observed in mice⁸, which we confirmed (Extended Data Fig. 1). In addition, we directly observed migration (Supplementary Videos 1, 2 and 3) and identified coexpression of integrin $\alpha_6\beta_4$ (Extended Data Figs 1 and 2). These cells also appear variably after bleomycin injury, in which approximately one-third of the Krt5⁺ cells resolved into type II pneumocytes by 50 days after injury (Extended Data Fig. 3). A cellular origin and mechanistic framework for expansion after influenza, and potential parallels in human lung injury, remain unknown.

To define the cell of origin, we lineage-traced mature cell types implicated in epithelial repair. Krt5⁺ cells appearing by day 11 after influenza infection were essentially completely untraced using Clara cell-specific protein (CC10) and surfactant protein C (SPC) Cre-recombinase drivers (CC10-CreERT2 and SPC-CreERT2, respectively, containing tamoxifen-inducible Cre-modified oestrogen receptor fusion proteins) (Fig. 1b–e and Extended Data Fig. 1i). Analysis at 7–8 days after injury confirmed mutual exclusivity of CC10-CreERT2-labelled and Krt5⁺ cells (Fig. 1b). Conflicting results in other reports are probably caused by tamoxifen persistence (Supplementary Discussion and Extended Data Fig. 4).

A small fraction (13%) of expanded Krt5⁺ cells bear the Krt5-CreERT2 lineage label (Fig. 1f, g), raising the possibility that tracheal basal cells might migrate distally during injury. We transplanted sections of fluorescent trachea into syngeneic animals and a non-fluorescent left lung into a fluorescent mouse⁹. Abundant Krt5⁺ cells arose after infection but none were fluorescent (Fig. 1h and Extended Data Fig. 1j, k). Upper-airway basal cells therefore do not contribute to this phenomenon and instead implicate a lineage-negative epithelial progenitor (LNEP) as the major source of Δ Np63⁺ and Krt5⁺ cells.

To characterize quiescent LNEPs, we used integrin β_4 expression in CC10-CreERT2 mice to segregate LNEPs from club cells in uninjured lungs (Fig. 2a) and confirmed minimal expression of mature lineage markers (Extended Data Fig. 5c). The CC10[−] β_4 ⁺ (LNEP-containing) population uniquely expressed Δ Np63 (Extended Data Fig. 5c). Δ Np63⁺ cells were identified *in situ* scattered sporadically throughout distal airways (Fig. 2c). These cells did not express detectable Krt5 protein (Extended Data Fig. 5a). In a total of 65 small airways examined in two mice, we identified 24 Δ Np63⁺ cells. Only 7 of the 24 cells were labelled in Krt5-CreERT2 mice (Fig. 2c and Extended Data Fig. 5a), probably explaining the small fraction of post-injury Krt5⁺ cells bearing the Krt5-CreERT2 lineage label (Fig. 1f, g).

Given the infrequency of Δ Np63⁺ cells, we suspected progenitor activity of the CC10[−] β_4 ⁺ population might be restricted to a smaller subset. Immunostaining revealed multicilia in 78% of this population, whereas Δ Np63⁺ cells were less than 1% (Extended Data Fig. 5b). To address this heterogeneity, we performed single-cell RNA sequencing (RNA-seq) on CC10[−] β_4 ⁺ cells and on rare Krt5-CreERT2-labelled cells, a subset of this population (Fig. 2i). The Δ Np63 transcript was detected in several cells in the CC10[−] β_4 ⁺ population (Fig. 2b, red circle, far left) as well as the Krt5-traced cells (Fig. 2b, green triangle). Analysis of variance (ANOVA) comparison between putative LNEPs (Krt5-traced cells combined with all p63-expressing cells) and the remaining cells revealed enrichment of ~900 genes (>2-fold change, >1 FPKM, $P \leq 0.05$) in the LNEP group (Supplementary Data 1). We note enrichment for pluripotency-associated transcription factors (Myc and Klf4) in the LNEP group (Fig. 2b), while many genes enriched in the remaining cells (Fig. 2b, top rows, Supplementary Data 2) are known markers of ciliated cells¹⁰. Surprisingly, Δ Np63⁺ CC10[−] β_4 ⁺ cells most closely related to the Krt5-traced cells also expressed cilia-associated genes (Fig. 2b, denoted by asterisk). Cytospins of CC10[−] β_4 ⁺ cells revealed primary cilia on Δ Np63⁺ cells and additional cells without discernible Δ Np63 (Fig. 2c, right), indicating that the LNEP profile extends to a larger fraction of Δ Np63-low or -negative cells.

To assess the potential of LNEPs *in vivo*, we devised a transplantation assay by which ~10⁵ fluorescent CC10[−] β_4 ⁺ cells were delivered orthotopically into influenza-injured mice (Fig. 2d). Seeded LNEPs developed into multicellular structures in two patterns seemingly dependent on location: areas of type II cells that were virtually indistinguishable

¹Department of Medicine, Cardiovascular Research Institute, University of California, San Francisco (UCSF), San Francisco, California 94143, USA. ²Department of Biochemistry, Stanford University School of Medicine and Howard Hughes Medical Institute, Stanford, California 94305, USA. ³Max Planck Institute for Evolutionary Anthropology, Department of Evolutionary Genetics, Deutscher Platz 6, 04103 Leipzig, Germany. ⁴Department of Anatomy, School of Medicine, University of California, San Francisco (UCSF), San Francisco, California 94143, USA.

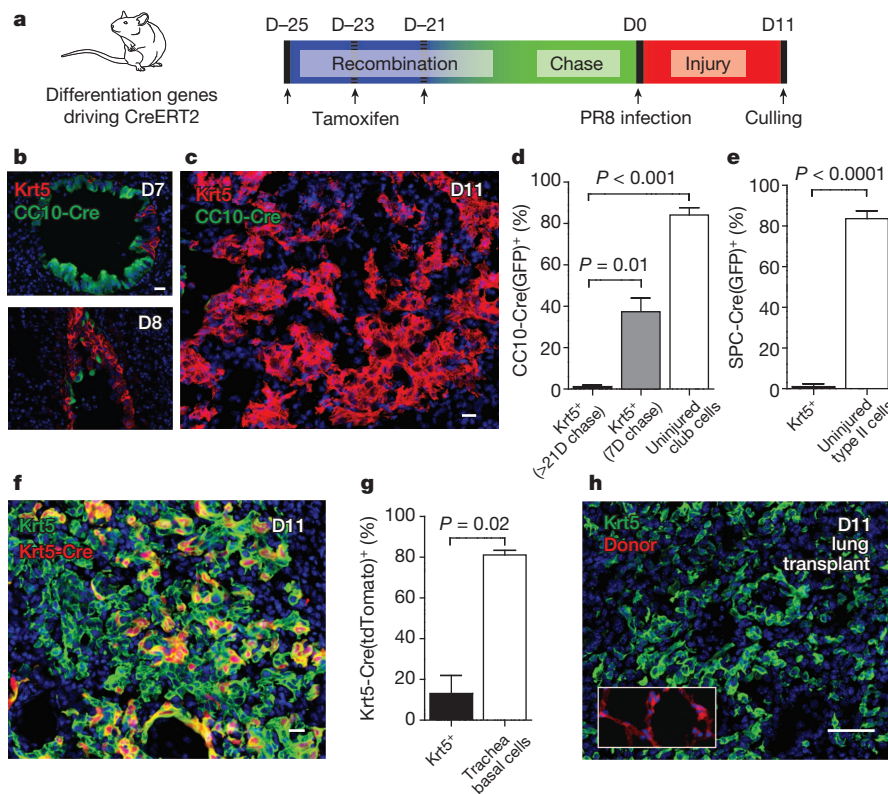


Figure 1 | Injury-induced Krt5⁺ cells are derived from a lineage-negative precursor. **a**, Schematic depicting lineage analysis methodology. **b**, **c**, Krt5⁺ cells are untraced (GFP-negative) after influenza injury in CC10-CreERT2/mTmG mice. **d**, **e**, Quantification of CC10 and SPC lineage tracing, expressed as a percentage of cells counted bearing the respective lineage tag (see Methods). Short chase time after tamoxifen administration to CC10-CreERT2 mice results in significant trace in Krt5⁺ cells (**e**) (Supplementary

Discussion). Data are mean \pm s.d., $n = 7$ CC10-CreERT2 and $n = 3$ SPC-CreERT2 mice quantified. **f**, **g**, A small fraction of Krt5⁺ cells bear the Krt5-CreERT2 trace (tdTomato⁺), quantified in **g** ($n = 3$ Krt5-CreERT2 mice). **h**, Krt5⁺ cells are not fluorescent after lung transplantation from a wild-type donor into a tdTomato recipient. Non-transplanted lung tissue retained fluorescence (inset). Image representative of one lung transplant. Scale bars, 20 μ m. Source data available online.

from surrounding endogenous type II cells (Extended Data Fig. 6a, b, h), and engraftments expressing Krt5 (Extended Data Fig. 6a, c) and CC10 (Extended Data Fig. 6g) near endogenous Krt5⁺ CC10⁺ structures. β_4^- type II cells engrafted infrequently in small clusters (<8 cells), and expressed only alveolar markers (Extended Data Fig. 6i). CC10⁺ cells could engraft but exhibited scant differentiation, even losing CC10 expression (Extended Data Fig. 6j, k). Transplantation of multi-ciliated cells resulted in only occasional persistence of single cells without structures (Extended Data Fig. 6l), consistent with their lack of progenitor properties^{11,12}.

Transplantation of mixed enhanced green fluorescent protein (eGFP)-labelled and tdTomato-expressing LNEPs demonstrated engraftments to be largely non-overlapping (Fig. 2e) and highly proliferative (Extended Data Fig. 6e), arguing for near-clonal expansion. Although mature type II cells do not express integrin β_4 (ref. 13), clones derived from donor LNEPs exhibited β_4 and SPC co-expression 5 days after transplant (Extended Data Fig. 6e), confirming their LNEP origin. These data demonstrate multipotency of LNEPs as well as the viability of orthotopic cell transplantation as a functional tool.

We interrogated the RNA-seq analysis and identified enrichment for CD14 in Δ Np63⁺ CC10⁻ β_4^+ cells (Fig. 2b, asterisk). In combination with CD200, which further selects against multi-ciliated cells (Extended Data Fig. 5e), CD14⁺ cells were isolated and transplanted. β_4^+ CD200⁺ CD14⁺ cells (~3,000) (Fig. 2f) phenocopied the larger (150,000) CC10⁻ β_4^+ population (Fig. 2g, h and Extended Data Fig. 7a–c), validating this small population as the active LNEPs. Using a complementary approach, distal Krt5-CreERT2-labelled Δ Np63⁺ cells within the LNEP fraction were transplanted (1,000 cells per mouse) (Fig. 2i). Multipotency was again observed, although we noted many fewer SPC-expressing cells (Fig. 2j, k and Extended Data Fig. 7d, e). We posit that

isolation using Krt5-driven Cre enriches for LNEPs that have undergone partial commitment to the Krt5 program, whereas surface-marker-based selection represents a less biased approach. This is consistent with lineage analysis (Fig. 1h) indicating Krt5-CreERT2-traced cells can only account for a small fraction of the Krt5⁺ expansion.

Accordingly, LNEPs cultured *ex vivo* did not express Krt5 even when treated with various trophic/morphogenic factors (Supplementary Table 2). However, bronchoalveolar lavage fluid (BALF) from injured mice induced marked proliferation and Krt5 expression. A total of $77 \pm 13\%$ (mean \pm s.d.) of colonies treated with the BALF stained positive for Krt5 (Fig. 3a–d), whereas type II cells treated with the same BALF did not respond.

Although the active principle(s) in injury BALF is uncertain, a screen of pathway inhibitors implicated a critical role of Notch. The γ -secretase inhibitor DAPT in conjunction with active BALF attenuated intensity and reduced the Krt5⁺ colony fraction ($22.7 \pm 13\%$) (Fig. 3c, d). This prompted us to analyse Notch activity *in vivo*. Notch1 intracellular domain (ICD) and the canonical Notch signalling target Hes1 were evident in the nucleus of parenchymal Krt5⁺ cells after influenza (Fig. 3e, f). Notch activity was further validated using a Notch reporter mouse (Cp-eGFP) (Extended Data Fig. 8a, b). When DAPT was administered to mice after influenza, the fraction of lung area bearing Krt5⁺ cells by day 11 was markedly reduced (Fig. 3g and Extended Data Fig. 7h).

During development, Notch signalling is known to suppress alveolar differentiation in both the lung and the mammary gland^{14,15}. When LNEPs were cultured in the presence of γ -secretase inhibitors, we observed strong induction of SPC expression, further promoted by 3-isobutyl-1-methylxanthine (IBMX)¹⁶ (Fig. 3h, i). Therefore, persistent Notch signalling prevents alveolar differentiation, whereas removal of this signal promotes maturation towards type II cells. This result proved

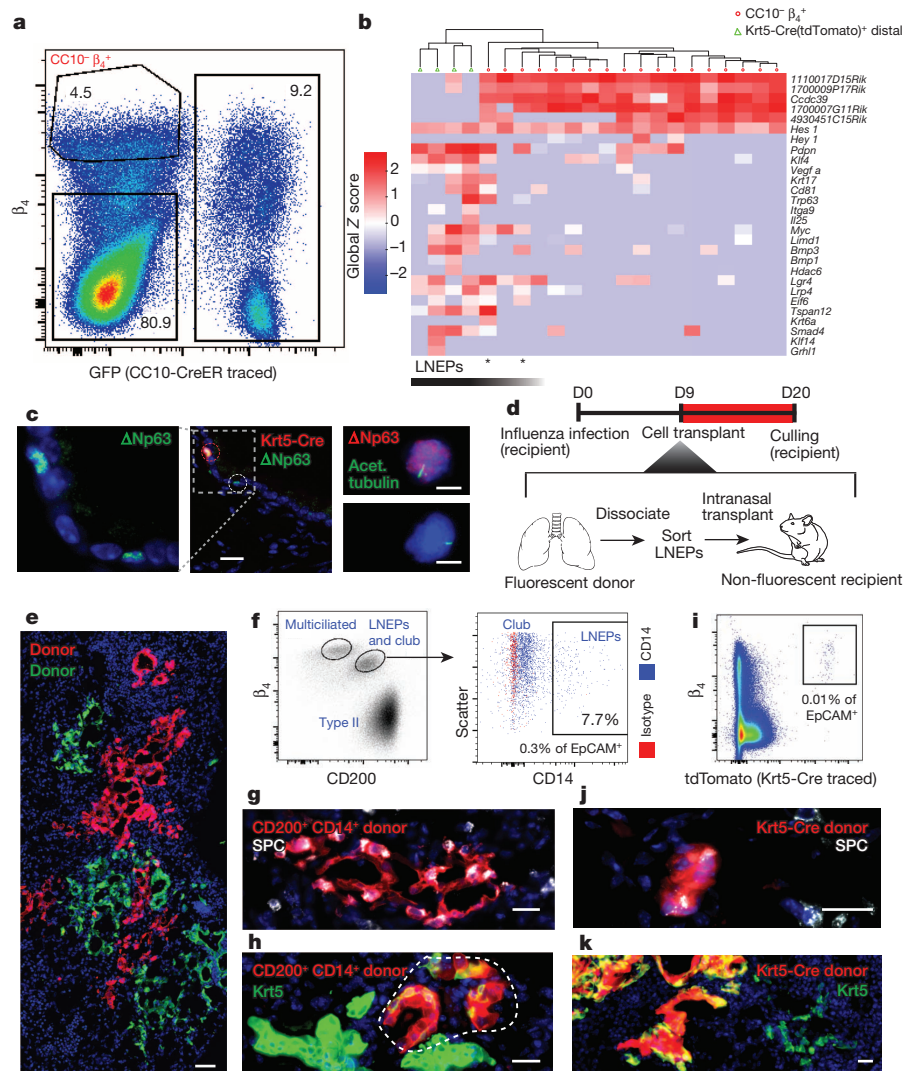


Figure 2 | Isolation and transplantation of a lineage-negative distal epithelial population. **a**, FACS segregation of epithelial (EpCAM⁺) cells by β₄ expression and a CC10-CreERT2 lineage tag (GFP), demonstrating a β₄⁺ population distinct from club cells. **b**, Hierarchical clustering/heat map of RNA-seq transcriptomes from single CC10⁺ β₄⁺ cells (circles) and distal Krt5-CreERT2-traced cells (triangles) (columns). Listed genes (rows) were selected from >1,200 differentially expressed genes identified by ANOVA. **c**, Immunofluorescent staining for ΔNp63 in uninjured lungs from Krt5-CreERT2 (tdTomato⁺) mice. Single cells from cytopspins of the CC10⁺ β₄⁺ population demonstrate primary cilium (green) in a subset of non-multiciliated

cells (right). **d**, Schematic depicting orthotopic cell transplantation methodology. **e**, Transplantation of LNEPs combined from eGFP- or tdTomato-expressing donors into a single recipient. Most engrafted regions are exclusively GFP⁺ (green) or tdTomato⁺ (red), suggesting clonal expansion. **f–h**, FACS isolation and transplantation of β₄⁺ CD200⁺ CD14⁺ LNEPs (f). Transplanted cells differentiate into both SPC⁺ (g) and Krt5⁺ (h) cells, representative of *n* = 3 transplants. **i–k**, FACS isolation and transplantation of Krt5-CreERT2-labelled LNEPs also differentiate into SPC⁺ (j) and Krt5⁺ (k) cells, representative of *n* = 2 transplants. Scale bars, 20 μm (c, left, g, h, j, k), 10 μm (c, right) and 100 μm (e).

relevant to the long-term outcome of regeneration in the influenza injury model.

Although regions of relatively normal histology bearing the Krt5-CreERT2 trace develop after resolution of bleomycin injury (Extended Data Fig. 3e), we were surprised to find few traced SPC⁺ type II cells after influenza (Extended Data Fig. 8e). Instead, large regions of Krt5-Cre-traced epithelial cysts were present in all mice examined between days 52 and 200 after injury (*n* = 7 mice) (Fig. 4a). These cysts consisted of CC10⁺ cells, scattered Krt5⁺ cells, otherwise nondescript epithelial cells, but very few SPC⁺ cells (Fig. 4a and Extended Data Fig. 8c–e), raising the possibility of ongoing Notch activity in cystic epithelium. Strong Hes1 expression persisted in Krt5-CreER-traced cyst-derived cells indefinitely (Fig. 4b and Extended Data Fig. 8f), whereas it is normally undetectable in alveolar epithelium (Extended Data Fig. 8g). The same correlation was observed in transplant experiments: LNEP-derived Krt5⁺ and CC10⁺ areas exhibited strong Hes1 expression, whereas it was low or absent in areas of type II cell differentiation (Extended Data

Fig. 6b–d). Notch antagonism *in vivo* via intranasal delivery of dibenzazepine (DBZ; in conjunction with dexamethasone and IBMX) resulted in a significant increase in the number of cyst-derived SPC⁺ cells (12.3% versus 1.6%) (Fig. 3j–l).

Persistent cysts bear a strong resemblance to micro-honeycombing in the lungs of patients with idiopathic pulmonary fibrosis (IPF) (Fig. 4c, d). These lungs (*n* = 10) showed almost all cystic epithelia were comprised of KRT5⁺ cells surrounded by either additional metaplastic KRT5⁺ cells or pseudostratified epithelium with ectopic but otherwise typical basal cells¹⁷. Distinct foci of hyperplastic SPC⁺ cells were also present. Notch activity correlated with KRT5⁺ cysts but was absent in most hyperplastic SPC⁺ cells (Fig. 4d–f and Extended Data Fig. 9a–d) and in normal alveolar regions (Extended Data Fig. 9k).

In lungs from patients with scleroderma (*n* = 7), fibrotic areas displayed the IPF pattern of persistent KRT5⁺ and HES1⁺ cystic structures (Extended Data Fig. 9e, g, h). However, in three less-fibrotic specimens, we observed extensive double-positive KRT5⁺ and SPC⁺ cells lining

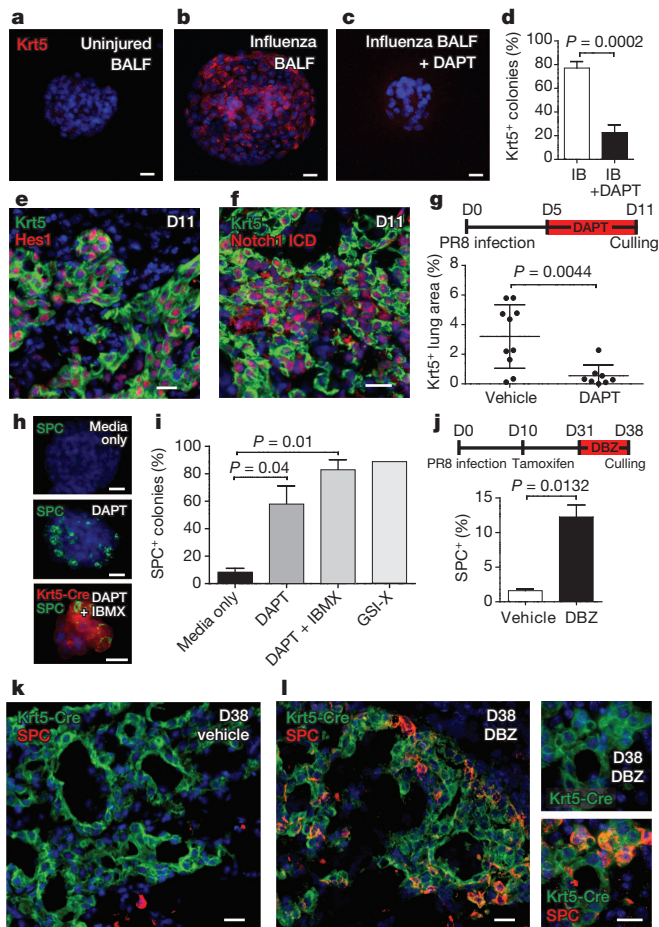


Figure 3 | Activation and Krt5 expression by lineage-negative progenitors is Notch-dependent. **a–d**, LNEP colonies upregulate Krt5 only after stimulation with BALF from influenza-injured mice (**b**) ($n = 6$ experiments), a process blocked by γ -secretase inhibition with DAPT (**c**) ($n = 4$ experiments), quantified in **d**. **e, f**, Hes1 (**e**) and Notch1 intracellular domain (ICD) (**f**) are present in the nucleus of Krt5⁺ cells at day 11 indicating Notch activity. **g**, γ -secretase inhibition during influenza injury reduces Krt5⁺ cell activation and expansion as measured by the fraction of lung section area. Each dot denotes one section; two sections per mouse, $n = 5$ (vehicle) or 4 (DAPT) mice per group. **h, i**, γ -secretase inhibition induces SPC expression in LNEPs *in vitro*, quantified in **i** ($n = 3$ experiments). **h**, Bottom, the Krt5-CreERT2 lineage label could be observed in SPC⁺ cells after DAPT treatment *in vitro*. **j–l**, Representative images of Notch inhibition *in vivo* via intranasal administration of DBZ results in a significant increase in Krt5-CreERT2-traced SPC⁺ cells (**l**) versus labelled cells in vehicle-treated mice (**k**) after influenza, quantified in **j** ($n = 2$ mice per group, >900 cells quantified per mouse in two sections from two separate lobes). IB denotes influenza BALF. Scale bars, 20 μ m. Data are mean \pm s.d. Source data for **g** and **j** are available online.

alveoli (Fig. 4g and Extended Data Fig. 9f). Although the origin of this KRT5⁺ expansion is uncertain in humans, we note Δ Np63⁺ KRT5⁺ cells in normal terminal airways (Extended Data Fig. 9i), analogous to LNEPs in mice.

These experiments identify a rare, undifferentiated epithelial population that is the major responder in distal lung after severe damage (Extended Data Fig. 10a). Notch signalling modulates the quiescence, activation and differentiation state of murine LNEPs (Extended Data Fig. 10b), providing a signalling model to frame the dynamic aspects of LNEP function. The persistently abnormal parenchymal structures that derive from LNEPs after influenza infection represent a failed regenerative process, promoted at least in part by ongoing Notch activity. The striking parallels to the currently inexplicable micro-honeycombing

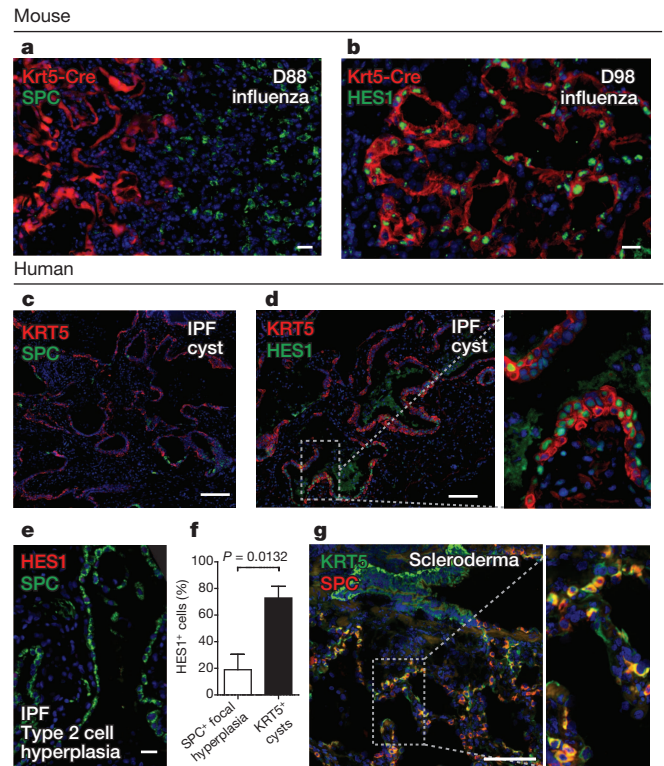


Figure 4 | Persistent Notch activity promotes cystic honeycombing in both mouse and human. **a**, Krt5-CreERT2-traced (tdTomato⁺) cells develop into cystic structures at late time points after influenza. **b**, Cyst cells demonstrate nuclear expression of Hes1 indicative of persistent Notch signalling. **c**, Lung from patient with IPF bearing honeycomb cysts with mutually exclusive KRT5⁺ and SPC⁺ cells. **d**, IPF honeycomb cysts with nuclear HES1 in KRT5⁺ cells and surrounding epithelium, similar to mouse (**b**). **e**, SPC⁺ type II cells in hyperplastic foci infrequently express HES1, quantified in **f** ($n = 8$ patients, mean \pm s.d.). **g**, Scleroderma lung demonstrating sub-pleural KRT5⁺ and SPC⁺ cell expansion with many KRT5⁺ SPC⁺ double-positive cells (right). Scale bars, 20 μ m (**a, b, d, e, g**) and 100 μ m (**c**).

that characterizes progressive fibrotic lung disease, including hyperactive Notch, suggest inappropriate Notch signalling may also be a major contributor to failed regeneration in chronic lung disease.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 March; accepted 27 November 2014.

Published online 24 December 2014.

- Clevers, H. The intestinal crypt, a prototype stem cell compartment. *Cell* **154**, 274–284 (2013).
- Gurtner, G. C., Werner, S., Barrandon, Y. & Longaker, M. T. Wound repair and regeneration. *Nature* **453**, 314–321 (2008).
- Tanimizu, N. & Mitaka, T. Re-evaluation of liver stem/progenitor cells. *Organogenesis* **10**, 208–215 (2014).
- King, R. S. & Newmark, P. A. The cell biology of regeneration. *J. Cell Biol.* **196**, 553–562 (2012).
- Desai, T. J., Brownfield, D. G. & Krasnow, M. A. Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* **507**, 190–194 (2014).
- Barkauskas, C. E. *et al.* Type 2 alveolar cells are stem cells in adult lung. *J. Clin. Invest.* **123**, 3025–3036 (2013).
- Giangreco, A. *et al.* Stem cells are dispensable for lung homeostasis but restore airways after injury. *Proc. Natl Acad. Sci. USA* **106**, 9286–9291 (2009).
- Kumar, P. A. *et al.* Distal airway stem cells yield alveoli *in vitro* and during lung regeneration following H1N1 influenza infection. *Cell* **147**, 525–538 (2011).
- Schrepfer, S. *et al.* Experimental orthotopic tracheal transplantation: the Stanford technique. *Microsurgery* **27**, 187–189 (2007).
- Treutlein, B. *et al.* Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **509**, 371–375 (2014).
- Rawlins, E. L., Ostrowski, L. E., Randell, S. H. & Hogan, B. L. M. Lung development and repair: Contribution of the ciliated lineage. *Proc. Natl Acad. Sci. USA* **104**, 410–417 (2007).

12. Rawlins, E. L. & Hogan, B. L. Ciliated epithelial cell lifespan in the mouse trachea and lung. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **295**, L231–L234 (2008).
13. Chapman, H. A. *et al.* Integrin $\alpha 6 \beta 4$ identifies an adult distal lung epithelial population with regenerative potential in mice. *J. Clin. Invest.* **121**, 2855–2862 (2011).
14. Guseh, J. S. *et al.* Notch signaling promotes airway mucous metaplasia and inhibits alveolar development. *Development* **136**, 1751–1759 (2009).
15. Chakrabarti, R. *et al.* E1f5 regulates mammary gland stem/progenitor cell fate by influencing notch signaling. *Stem Cells* **30**, 1496–1508 (2012).
16. Wang, J. *et al.* Differentiated human alveolar epithelial cells and reversibility of their phenotype *in vitro*. *Am. J. Respir. Cell Mol. Biol.* **36**, 661–668 (2007).
17. Seibold, M. A. *et al.* The idiopathic pulmonary fibrosis honeycomb cyst contains a mucociliary pseudostratified epithelium. *PLoS ONE* **8**, e58658 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by National Institutes of Health (NIH) grants R01 HL44712 and U01 HL111054 and a sponsored research agreement with Daiichi Pharmaceuticals. A.E.V. is supported by F32 HL117600-01. The authors thank T. Kim for assistance with animal work and K. Corbin for assistance with Imaris software. Mouse line art was created by A. van de Wiel. We also thank P. Wolters at the UCSF Interstitial Lung Disease Blood and Tissue Repository for procuring diseased lung

tissues. We thank the Nina Ireland Program for Lung Health that supported the tracheal/lung transplant experiments.

Author Contributions A.E.V. and H.A.C. designed the study, analysed the data, and wrote the manuscript. A.E.V. performed lineage tracing, flow cytometry purification, and characterization of lung cells; J.E.G. titred PR8 virus and initiated all infections; A.N.B. isolated lung cell suspensions, assisted with flow cytometry, and designed and performed most of the immunostaining; Y.X. assisted with biochemistry, RNA analysis, and immunostaining; K.T. and V.T. managed the mouse genotyping and performed *in vivo* mouse experiments; V.T. isolated lung cells and designed quantification methods; F.C.L. and M.R.L. performed lung transplantations; M.M. procured and screened human lungs; D.G.B. and B.T. synthesized libraries and provided initial data analysis for RNA-seq experiments J.R.R. provided key reagents and assisted with study design.

Author Information The FPKM files from single-cell RNA-seq experiments have been deposited in Gene Expression Omnibus (GEO) under accession GSE61300. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.E.V. (andrew.vaughan@ucsf.edu) or H.A.C. (hal.chapman@ucsf.edu).

METHODS

Animals. SPC-CreERT2 ($\text{Sftpc}^{\text{tm1}(\text{cre/ERT2}, \text{rtTA})\text{Hap}}$), Krt5-CreERT2 ($\text{Krt5}^{\text{tm1}(\text{cre/ERT2})\text{Blh}}$), CC10-CreERT2 ($\text{Scgb1a1}^{\text{tm1}(\text{cre/ERT2})\text{Blh}}$), FoxJ1-CreERT2 ($\text{Tg}(\text{Foxj1-cre/ERT2})\text{1Blh}$) and Cp-eGFP ($\text{Tg}(\text{Cp-EGFP})\text{25Gaia}$) mice are previously described^{13,18–20}. All of these strains were bred to either mTmG ($\text{Gt}(\text{ROSA})\text{26Sor}^{\text{tm4}(\text{ACTB-tdTomato}, -\text{EGFP})\text{Luo}}$)²¹ or Ai14-tdTomato ($\text{Gt}(\text{ROSA})\text{26Sor}^{\text{tm14}(\text{CAG-tdTomato})\text{Hze}}$)²² mice to generate mice expressing a fluorophore in Cre-expressing cells. SPC-CreERT2, Krt5-CreERT2 and CC10-CreERT2 mice were developed in a 129 background and backcrossed to C57BL6 for at least three generations. For transplant experiments, mTmG and/or Ub-GFP ($\text{Tg}(\text{UBC-GFP})\text{30Scha}$)²³ were used for donor cells. For all experiments, 6–8-week-old animals of both sexes were used in equal proportions. Investigators were not blinded to mouse identity. All studies were approved by University of California, San Francisco (UCSF) Institutional Animal Care and Use Committees (IACUC), protocol AN088356-03. All animal studies used a minimum of three mice per group with the exception of DBZ (see below).

For lineage analysis of the cell of origin of Krt5⁺ cells, mice were administered three doses (Krt5-CreERT2) or five doses (SPC-CreERT2 and CC10-CreERT2) of 0.25 mg g^{−1} body weight tamoxifen in 50 μ l corn oil. A chase period of >21 days was used to insure the absence of residual tamoxifen before injury.

Injury (influenza, bleomycin). Mice were administered 280 focus-forming units (FFU) of influenza A/H1N1/Puerto Rico/8/34 (PR8) intranasally. PR8 virus dissolved in 30 μ l PBS was pipetted onto the nostrils of heavily anaesthetized mice (visual confirmation of agonal breathing), whereupon mice aspirated the fluid directly into their lungs. The mice were allowed to recover and weighed twice a week. For experiments analysing the lineage fate of Krt5⁺ cells, a single dose of 0.125 mg g^{−1} body weight tamoxifen was administered at day 10 after PR8 infection.

Infective viral particles were assayed by inoculation of either stock virus or homogenate (in 1 ml PBS) of left lung, spleen and brain onto 96-well plates of confluent MDCK cells. After 1 h, samples were decanted and replaced with serum-free media containing modified trypsin (1-(tosylamido-2-phenyl) ethyl chloromethyl ketone (TPCK)-treated) at 100 μ g ml^{−1}. Fifteen hours later, the cells were fixed in 100% methanol, and then subjected to indirect immunocytochemistry using Millipore mouse anti-influenza A (MAB 8257) at 1.25 μ g ml^{−1}, followed by Vector 102 biotinylated horse anti-mouse, and the biotin/avidin system (PK-4002) with diaminobenzidine as a chromogen. Samples were processed in triplicate over dilutions, and foci were counted in wells that yielded 30–100 discrete foci.

Bleomycin (1.7 U kg^{−1} body weight) was administered intratracheally. Mice were weighed twice a week. For lineage tracing Krt5⁺ cells after bleomycin, a single dose of 0.125 mg g^{−1} body weight tamoxifen was administered at day 17 after bleomycin.

Treatment of animals with γ -secretase inhibitors. For DAPT administration, mice received 50 mg kg^{−1} body weight DAPT in 20 μ l dimethylsulphoxide (DMSO) per intraperitoneal injection, for the indicated periods. For intranasal DBZ administration, 30 μ mol kg^{−1} body weight DBZ was suspended in 50 μ l sterile PBS and sonicated in a Bioruptor UCD-200 for 15 min total, 30 s intervals on ice. In DBZ experiments, both DBZ and vehicle group also received 2.5 μ g g^{−1} body weight dexamethasone (Sigma) in the intranasal solution and 10 mg kg^{−1} IBMX (Sigma) intraperitoneally daily. Both DAPT and DBZ were obtained from Toronto Research Chemicals. More than 7,500 Krt5-CreERT2-labelled cells were quantified for SPC expression in at least two individual lobes from each mouse.

Orthotopic lung transplantation. Left lung transplants were carried out using the method described previously²⁴. The donor animal was anaesthetized and injected with heparin (50 U) immediately before perfusion of the lung vasculature with 5 ml of ice-cold Perfadex solution (Xvivo Perfusion), clamping of the hilar structures, and removal from the donor animal. The left lung was transplanted into the recipient animal using the cuff anastomosis technique.

Orthotopic tracheal transplantation. The donor animal was anaesthetized and with the aid of microscopic dissection, a segment of trachea composed of 5–7 tracheal rings was removed. The recipient animal was anaesthetized and the donor trachea was interposed using proximal and distal anastomoses⁹.

Immunofluorescence analysis of tissue. After euthanasia, lungs were either immediately inflated with OCT and flash frozen or inflated with 4% paraformaldehyde (PFA) and fixed for 1 h at room temperature and subsequently embedded in OCT. Sections (7 μ m) were cut on a cryostat, with fresh-frozen tissue immediately fixed for 5 min in 4% paraformaldehyde at room temperature. All sections were subsequently incubated for 3 \times 10-min intervals with 1 mg ml^{−1} sodium borohydride (Sigma) in PBS to reduce aldehyde-induced background fluorescence. Slides were subsequently blocked \geq 1 h in PBS plus 1% bovine serum albumin (Affymetrix), 5% nonimmune horse serum (UCSF Cell Culture Facility), 0.1% Triton X-100 (Sigma) and 0.02% sodium azide (Sigma). Slides were incubated overnight in primary antibodies listed below, diluted in block solution. Slides were washed three times with PBS plus 0.1% Tween 20, and incubated with secondary antibodies (typically Alexa Fluor conjugates, Life Sciences) at a 1:2,000 dilution \geq 1 h. Finally, slides were again

washed, incubated with 1 μ M DAPI for 5 min, and mounted using Prolong Gold (Life Sciences).

The following antibodies were used: rabbit anti-proSPC (1:3,000; Millipore, AB3786), goat anti-proSPC (1:2,000; Santa Cruz, M-20), goat anti-CC10 (1:10,000, a gift from B. Stripp), rabbit anti-Krt5 (1:1,000; Covance, PRB-160P), chicken anti-Krt5 (1:1,000; Covance, SIG-3475), rabbit anti-ANp63 (1:100; Biogenex, POLY6190), rat anti-CD45 (1:200, BD 30-F11), sheep anti-eGFP (1:500; Pierce, 10396164), rabbit anti-phospho histone H3 (1:500; Millipore, 06-570), rabbit anti-Hes1 (1:1,000; Cell Signaling, D6P2U), rabbit anti-activated Notch1 (1:1,000; Abcam, ab8925), mouse anti-acetylated tubulin (1:500, Sigma, 6-11B-1).

Quantification of lineage tracing. Samples were prepared for immunofluorescence staining. Quantification at day 11 after influenza is the result of counting >2,900 cells (CC10 trace), >4,000 cells (SPC trace), or >1,300 (Krt5 trace) from at least three mice per genotype. Cells were counted from over five sections per mouse and included at least three individual lobes. Mutual exclusivity of CC10-traced and Krt5⁺ cells at days 7–8 was determined with a smaller sample size, $n = 2$ mice, 12 Krt5⁺ airways, >500 cells examined. Only mice possessing the appropriate genotype were used in studies.

Epithelial cell isolation and flow cytometry. Lung epithelial cells were isolated as previously described¹³, with the following modifications. After installation with agarose and subsequent hardening by a brief incubation on ice, each lobe was cut away from the mainstem bronchi. The proximal-most quarter of each lobe surrounding the bronchi was then cut away to minimize the inclusion of basal cells in the cell preparation, and the previous protocol was followed from this point on.

For FACS analysis, single-cell preparations were incubated for 30–45 min at 4 °C with the following primary antibodies: phycoerythrin (PE), Alexa Fluor 488, or BV421-conjugated rat anti-mouse EpCAM (1:500; Biogenex, G8.8), Alexa Fluor 647 or PE-conjugated rat anti-mouse integrin β_4 (1:75; BD, 450-9D), Alexa 647-conjugated CD200 (1:100, Biogenex, OX-90), and PE/Cy7-conjugated CD14 (1:100, Biogenex, Sa14-2). Antibody incubations were done in DMEM (without phenol red) plus 2% FBS, and cells were washed twice with PBS after antibody incubations. Sorting and analysis was performed on BD FACS Aria cytometers.

Orthotopic cell transplantation. Recipient C57BL/6 mice were infected with PR8 (see Animals). At 9 days after infection, donor cells were sorted from mTmG or Ub-GFP mice (Animals) and resuspended in 50 μ l sterile PBS. Recipient mice received cell solution intranasally as described above for influenza administration. The total number of β_4 ⁺ cells ranged from 150,000 to 350,000 per transplant ($n = 6$), and equivalent numbers of β_4 [−] cells were always transplanted into injured littermates for comparison. For transplantation of Krt5-CreERT2-labelled cells, 1,000 cells were transplanted per recipient ($n = 2$). For β_4 ⁺ CD14⁺ CD200⁺ cell transplants, 3,000–10,000 cells were transplanted per mouse ($n = 3$). FoxJ1-CreERT2-labelled or CC10-CreERT2-labelled cell transplants were performed in $n = 3$ or 4 mice each, respectively (1×10^5 – 3×10^5 cells per mouse). Endpoint analysis was performed at day 21 after infection unless otherwise noted. For analysis of proliferation, recipient mice were administered 50 mg kg^{−1} body weight Edu (Santa Cruz) in PBS daily. Edu was detected with Click-iT Edu Alexa Fluor 488 Imaging Kit (Invitrogen).

Primary cell culture. Isolated primary lung epithelial cells were plated and cultured on Matrigel as follows. Eight-well chamber slides were coated with 150 μ l Matrigel per well, allowed to solidify at 37 °C, and then equilibrated with SABM (Lonza) for at least 30 min before cell plating. A total of 15,000–40,000 cells were plated in each well of and maintained in 'baseline' media consisting of SAGM (Lonza) supplemented with 5% charcoal-stripped FBS and 10 ng ml^{−1} KGF (FGF-7, Peprotech). Other growth factors were included in the media only when indicated and are summarized in Supplementary Table 2.

BALF was collected from injured animals for cell culture as follows. Euthanized mice were intratracheally intubated before cardiac perfusion and 1 ml of baseline media was lavaged. The lungs were repeatedly lavaged with the media at least three times. BALF was then centrifuged three times for 5-min spins at 1,500g to remove the cells and other debris. Clarified BALF was finally filtered through a 0.25- μ m Spin-X filter (Sigma) to remove any additional debris and to ensure a cell-free preparation. BALF prepared in this way was either added to cells immediately or frozen in aliquots at −80 °C and added to cultured cells without dilution.

Long-term cell culture. Cells isolated as above were maintained in SAGM as above, with the addition of 10 μ M Y-27632 (Sigma) and 50 ng ml^{−1} murine noggin (Peprotech). Cells were passaged every 7–10 days by initial incubation with 25 U ml^{−1} dispase at 37 °C for 20 min to liberate colonies. Single-cell dissociation was performed by additional 10-min incubation with 2 mM EDTA in PBS in combination with mechanical disaggregation by pipetting.

γ -secretase treatment of LNEPs *in vitro*. LNEPs maintained as above were dissociated and re-plated directly into SAGM baseline media with added DAPT or GSI-X (Calbiochem) at 40 or 20 μ M concentrations (unless otherwise indicated).

For SPC induction experiments, IBMX was added when indicated. LNEPs were cultured for 7–10 days and then analysed by immunofluorescent staining.

Immunofluorescence analysis of cultured cells. Cells grown on matrigel were fixed for 5–10 min in IHC Zinc Fixative (BD) and subsequently stained as indicated above, except that all staining solutions were prepared with TBS as the zinc fixative reacts with phosphate.

Live slice imaging. Krt5-CreERT2/TdTomato mice were administered 280 FFU PR8 (as above) and received a single 0.25 mg kg^{-1} dose of tamoxifen 24 h before culling at the indicated time points. Injured mice were euthanized and perfused and lavaged with PBS. Lungs were instilled with 2% low-melting point agarose and $\sim 300\text{-}\mu\text{m}$ slices were prepared on a vibratome. Lung slices were maintained in SAGM plus 10 ng ml^{-1} KGF during imaging with the addition of 500 nM hydroxytamoxifen (Sigma) to induce recombination in all Krt5-expressing cells. Slices were imaged continuously for 12 h in a 37°C chamber on an inverted stage with a Leica SP5 confocal microscope. Images obtained were deconvoluted with Bitplane Imaris for presentation.

Quantitative reverse transcriptase PCR. RNA was isolated from sorted cells using the Promega RNA Reliprep kit. cDNA was synthesized and amplified using the Ovation PicoSL WTA V2 kit (NuGen). Reverse transcription PCR (RT-PCR) reactions were performed using Faststart Universal SYBR green Master Mix (Roche) and run on an Eppendorf Realplex² thermocycler. Primer sequences are as listed: SPC (also known as *Sftpc*), forward, 5'-ATGGACATGAGTAGCAAAGAGGT-3', reverse 5'-CACGATGAGAAGGCGTTTGAG-3'; CC10 (also known as *Scgb1a1*), forward, 5'-ATGAAGATCGCCATCACAATCAC-3', reverse 5'-GGATGCCACATAACCAGACTCT-3'; Krt5 forward, 5'-TCCAGTGTGTCCTCCGAAGT-3', reverse 5'-TGCCTCCGCCAGAACTGTA-3'; ΔNp63 , forward, 5'-ATGTTGTA CCTGGAAAACAATGCC-3', reverse, 5'-CAGGCATGGCACGGATAAC-3'; Jag1 forward, 5'-CCTCGGGTCAGTTTGAGCTG-3', reverse, 5'-CCTTGAGGCACACTTTGAAGTA-3'; Jag2, forward, 5'-CAATGACACCACTCCAGATGAG-3', reverse, 5'-GGCCAAAGAAGTCGTTGCG-3'; Hey1 forward, 5'-GCGCGGACGAGAATGGAAAA-3', reverse, 5'-TCAGGTGATCCACAGTCATCTG-3'; Aqp5, forward, 5'-AGAAGGAGGTGTTCAGTTGC-3', reverse, 5'-GCCAGAGTATGGCCGGAT-3'; Abca3 forward, 5'-GCTTGAAGATCCAGTCGGAGA-3', reverse 5'-CATAGCGAATGTAGTCCTCGAAG-3'; Cgrrp (also known as *Calca*), forward, 5'-GCGGGCTCTAGCTTGGACAG-3', reverse, 5'-AAGGTGTGAAA CTTGTTGAGGT-3'; Sox2, forward, 5'-GCGGAGTGGAAACTTTTGTTC-3', reverse, 5'-GGGAAGCGTGACTTATCCTTCT-3'; Foxj1, forward, 5'-CATC

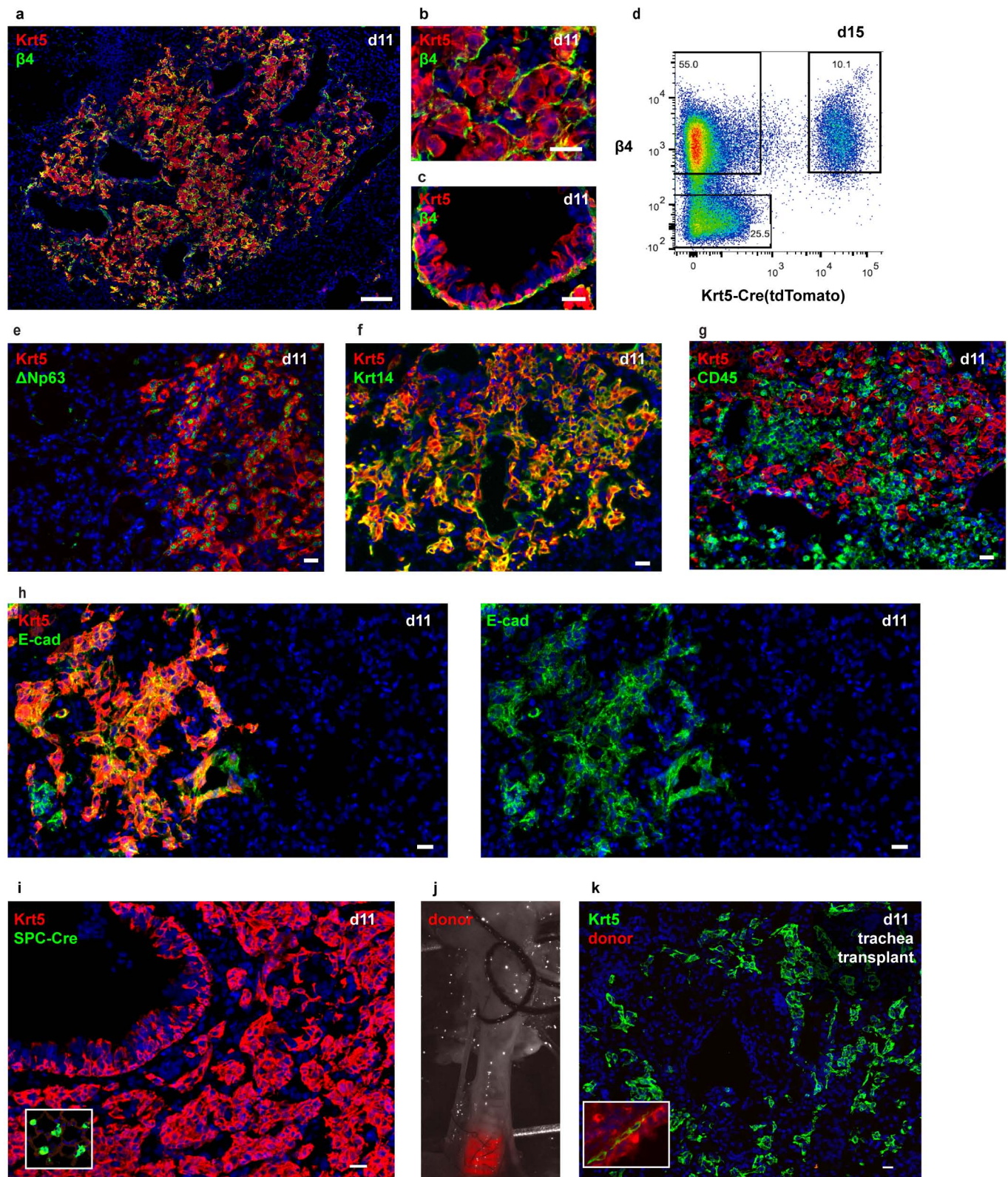
TACAAGTGGATCACGGAC-3', reverse, 5'-GAGCAGGCGCTCTGCGTAC TG-3'.

Single-cell RNA-seq. Distal lung epithelial cells were isolated and FACS sorted as described above from CC10-CreERT2/mTmG mice. In addition, tdTomato⁺ cells were sorted from tamoxifen-treated Krt5-CreERT2/TdTomato mice. Sorted single cells were captured on a medium-sized ($10\text{--}17 \mu\text{m}$ cell diameter) microfluidic RNA-seq chip (Fluidigm) using the Fluidigm C1 system. All downstream steps (lysis, cDNA synthesis/amplification, library preparation, sequencing and raw data processing) were performed exactly as previously described¹⁰. Fragment per kilobase of exon per million fragments mapped (FPKM) files for each cell were analysed using Fluidigm Singular software running in R.

Human tissues. All human tissue samples were obtained from UCSF Interstitial Lung Disease Blood and Tissue Repository and are classified as Non-identifiable Otherwise Discarded Human Tissues.

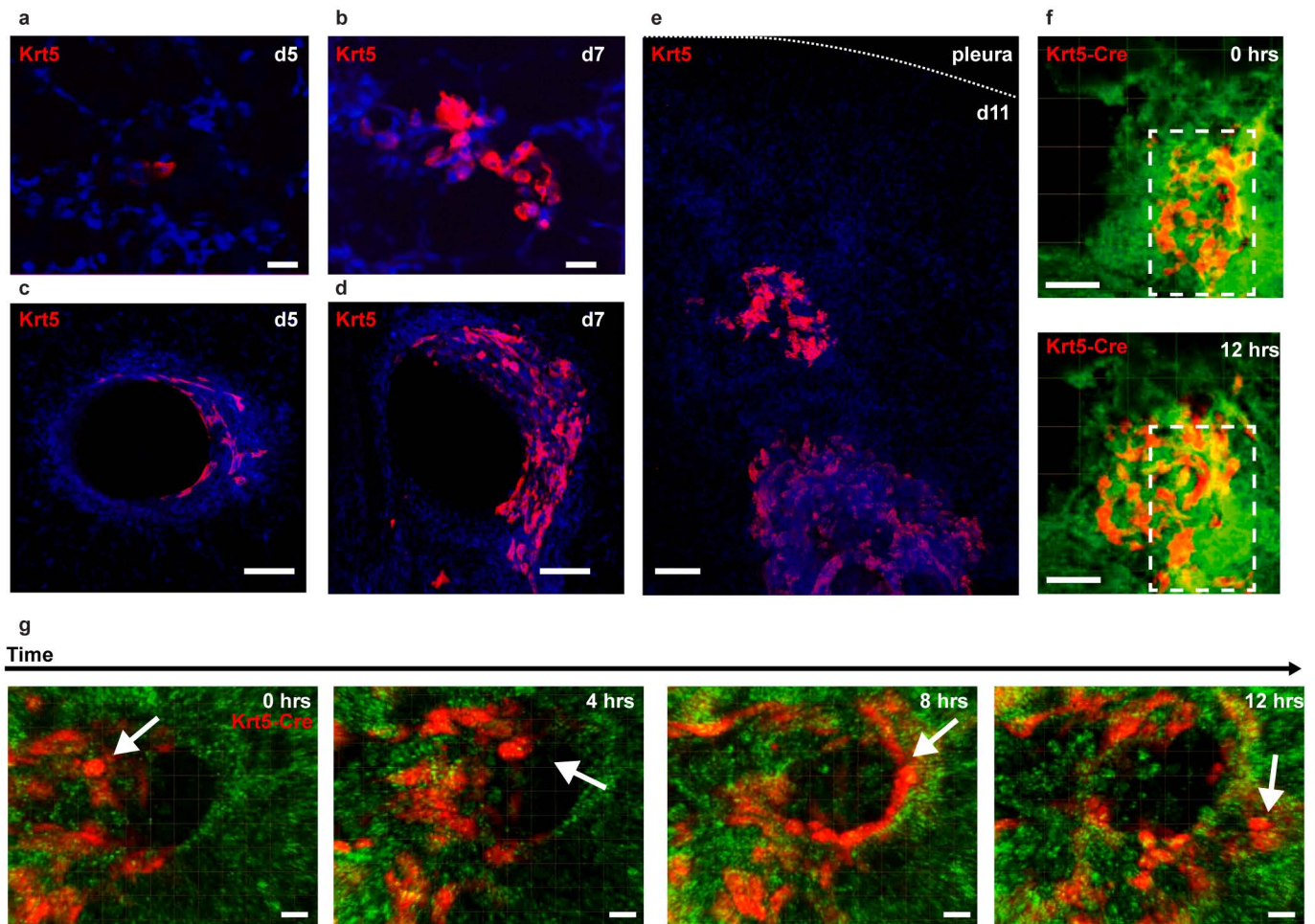
Statistics. For calculations involving single cell RNA-seq, Fluidigm Singular software running in R was used. All other statistical calculations were performed using Graphpad Prism. P values were calculated from two-tailed t-tests (paired or unpaired depending on experimental design) or ANOVA for multivariate comparisons. Variance was analysed at the time of t-test analysis. This data are not included in the manuscript but is available upon request. No statistical method was used to pre-determine sample size.

18. Van Keymeulen, A. *et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479**, 189–193 (2011).
19. Rawlins, E. L. *et al.* The role of Scgb1a1⁺ Clara cells in the long-term maintenance and repair of lung airway, but not alveolar, epithelium. *Cell Stem Cell* **4**, 525–534 (2009).
20. Duncan, A. W. *et al.* Integration of Notch and Wnt signaling in hematopoietic stem cell maintenance. *Nature Immunol.* **6**, 314–322 (2005).
21. Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. A global double-fluorescent Cre reporter mouse. *Genesis* **45**, 593–605 (2007).
22. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neurosci.* **13**, 133–140 (2010).
23. Schaefer, B. C., Schaefer, M. L., Kappler, J. W., Marrack, P. & Kedl, R. M. Observation of antigen-dependent CD8⁺ T-cell/ dendritic cell interactions *in vivo*. *Cell Immunol.* **214**, 110–122 (2001).
24. Krupnick, A. S. *et al.* Orthotopic mouse lung transplantation as experimental methodology to study transplant and tumor biology. *Nature Protocols* **4**, 86–93 (2009).



Extended Data Figure 1 | Characterization of influenza-induced $Krt5^{+}$ cells. a–c, Alveolar (a, b) and airway (c) $Krt5^{+}$ cells strongly express $\beta4$ after influenza injury. d, FACS plot of epithelial (EpCAM $^{+}$) cells from tamoxifen-treated $Krt5$ -CreERT2/ $tdTomato$ mice at day 15 after influenza, demonstrating $\beta4$ expression in nearly all traced ($tdTomato^{+}$) cells. e, f, Most $Krt5^{+}$ cells co-express $\Delta Np63$ (e) and $Krt14$ (f). g, h, Expanded $Krt5^{+}$ cells are invariably associated with abundant CD45 $^{+}$ inflammatory cells (g) and few if any

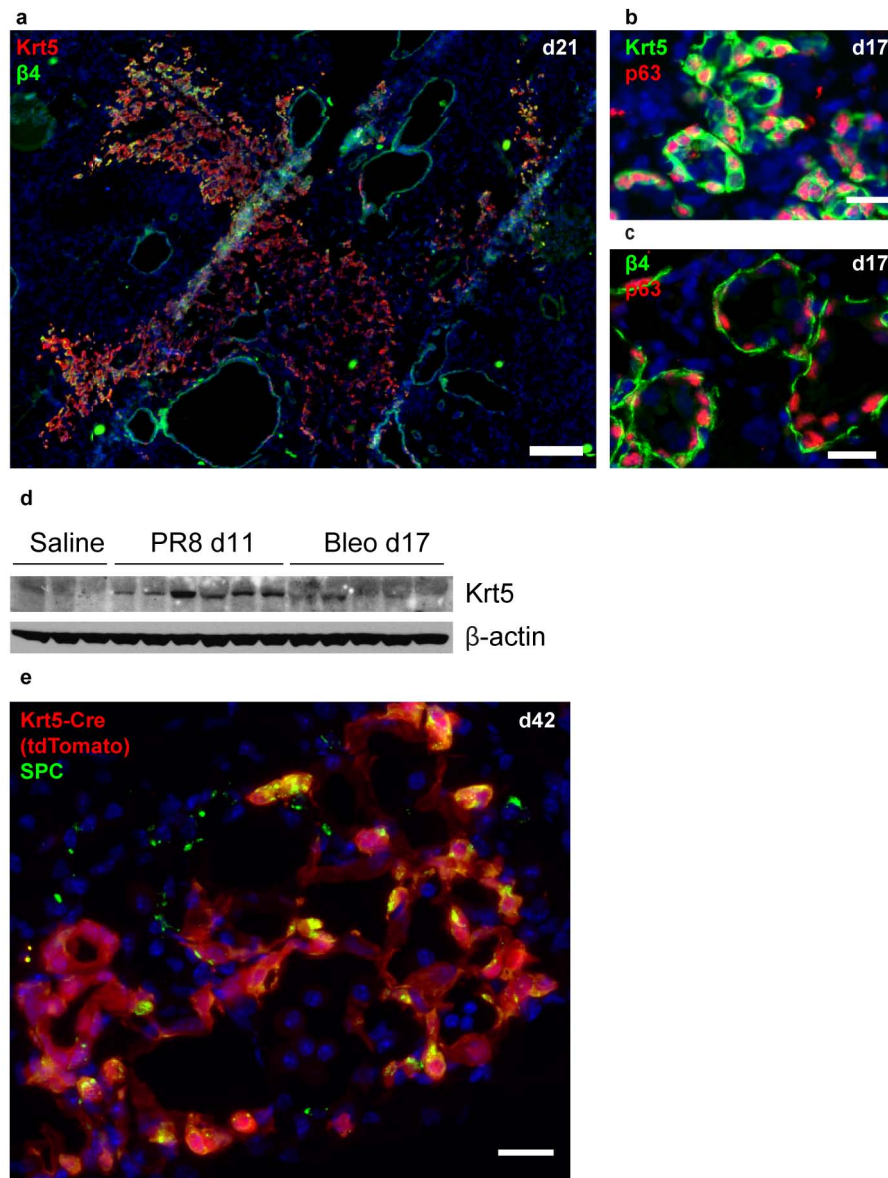
remaining normal E-cadherin $^{+}$ epithelial cells other than the $Krt5^{+}$ cells themselves (h). i, $Krt5^{+}$ cells are unlabelled in SPC-CreERT2/ $mTmG$ mice. Inset in i demonstrates appropriate labelling of type II cells in an uninjured region of the same lung. j, k, $Krt5^{+}$ cells are not fluorescent after trachea transplantation from $tdTomato$ donor. Basal cells in transplanted section of trachea retained fluorescence (j, inset in k). Scale bars, 100 μm (a) and 20 μm (b, c, e–k).



Extended Data Figure 2 | Influenza-induced Krt5⁺ cells arise in both airways and alveoli and migrate across, around and through airway and parenchymal tissue. **a, b,** Krt5⁺ cells are detected in alveoli as early as day 5 and are found in larger clusters over time. **c, d,** Krt5⁺ cells similarly arise in airways in greater abundance with time. **e,** Distinct alveolar and airway expansion is apparent 11 days after infection. **f,** Freeze-frames of live imaging

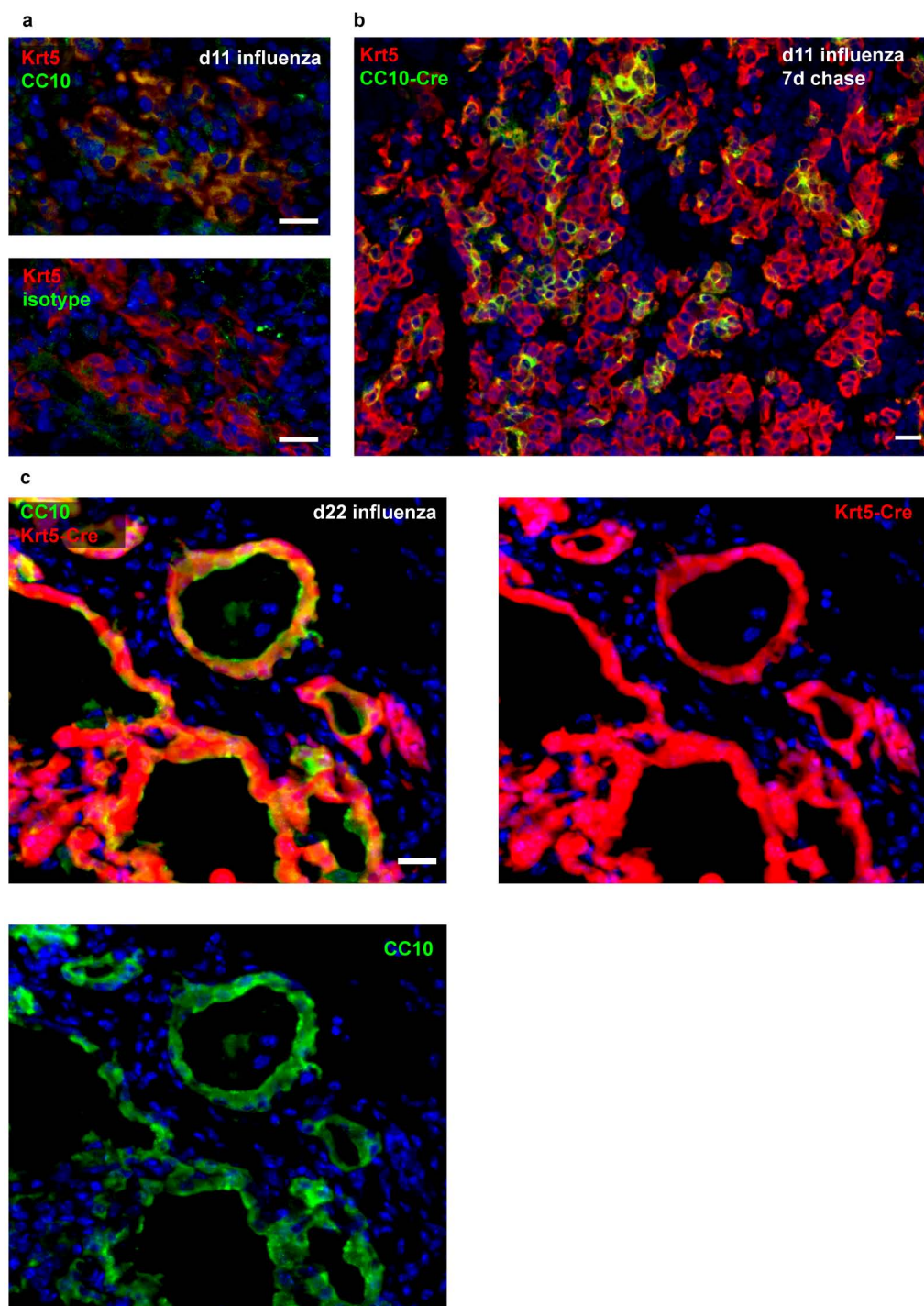
from a Krt5-CreERT2/tdTomato mouse 11 days after influenza, in which tdTomato⁺ cells migrate from their original location (white box) outward. See Supplementary Video 1. **g,** Freeze-frames from a small airway in the same mouse; arrow denotes a single cell crossing the basement membrane. See Supplementary Video 2. Scale bars, 20 μ m (**a, b, g**) and 100 μ m (**c, d, f**).

bleomycin



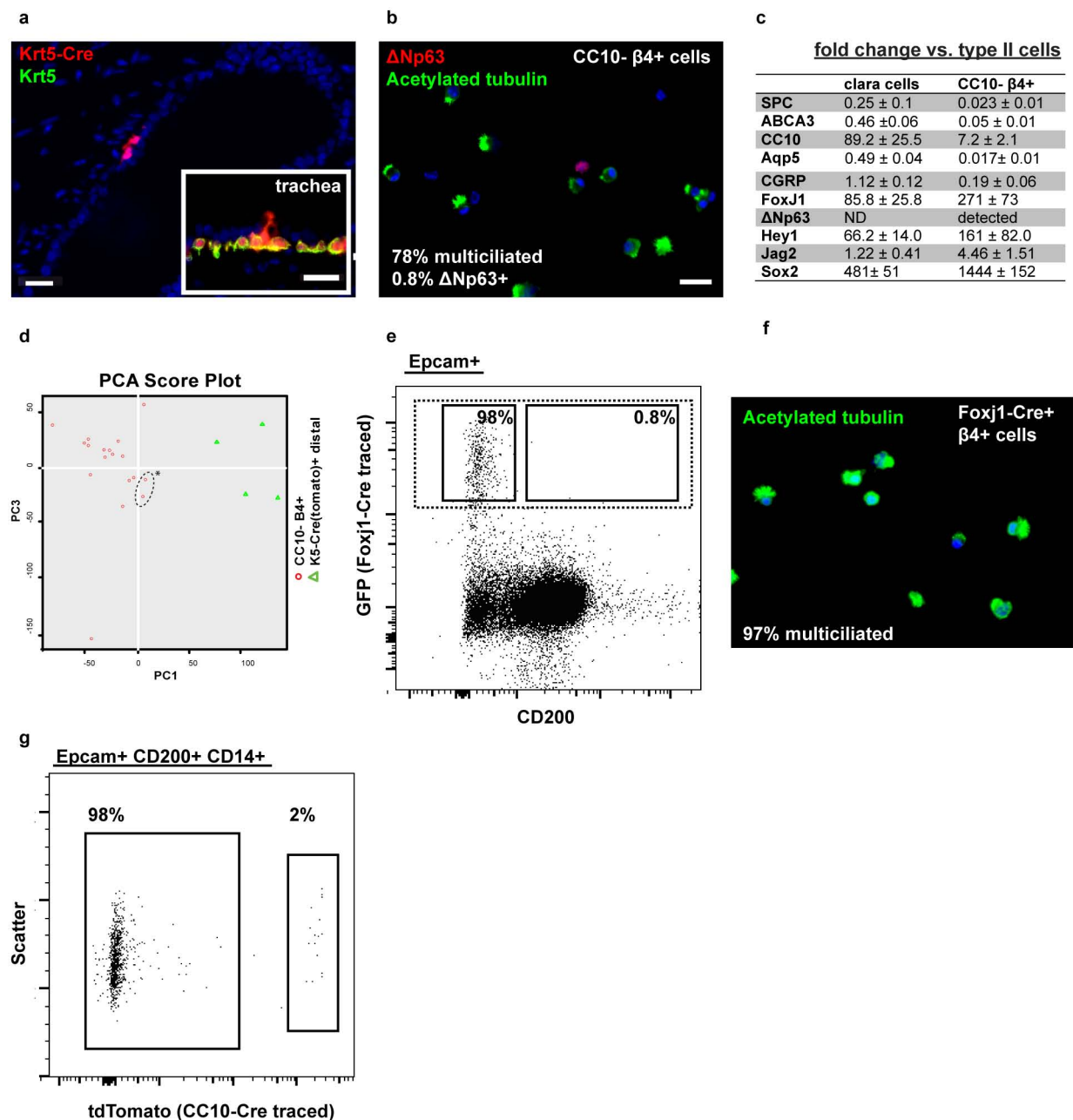
Extended Data Figure 3 | Characterization of bleomycin-induced Krt5⁺ cells. a–c, β₄⁺ Krt5⁺ cells also arise after bleomycin injury and express ΔNp63 (b, c). d, Western blotting demonstrating more pronounced and reproducible Krt5 induction after influenza injury at day 11 than after bleomycin injury at day 17. Each lane was loaded with whole-lung lysate from a single mouse; average percentage lung area corresponding to a band in influenza-injured mice is 3.6 ± 0.5% (*n* = 13 mice quantified, see Fig. 3g as an example).

e, Lineage tracing of bleomycin-injured Krt5-CreERT2 mice reveal traced (tdTomato⁺) type II cells expressing SPC and cells morphologically resembling type I cells. In total, 31% of Krt5-CreERT2 traced cells express SPC by day 50 after bleomycin (*n* = 3 mice, 264 Krt5-CreERT2-labelled cells counted). Scale bars, 100 μm (a) and 20 μm (b, c, e). Full western blot scan in d is available as Supplementary Fig. 1.



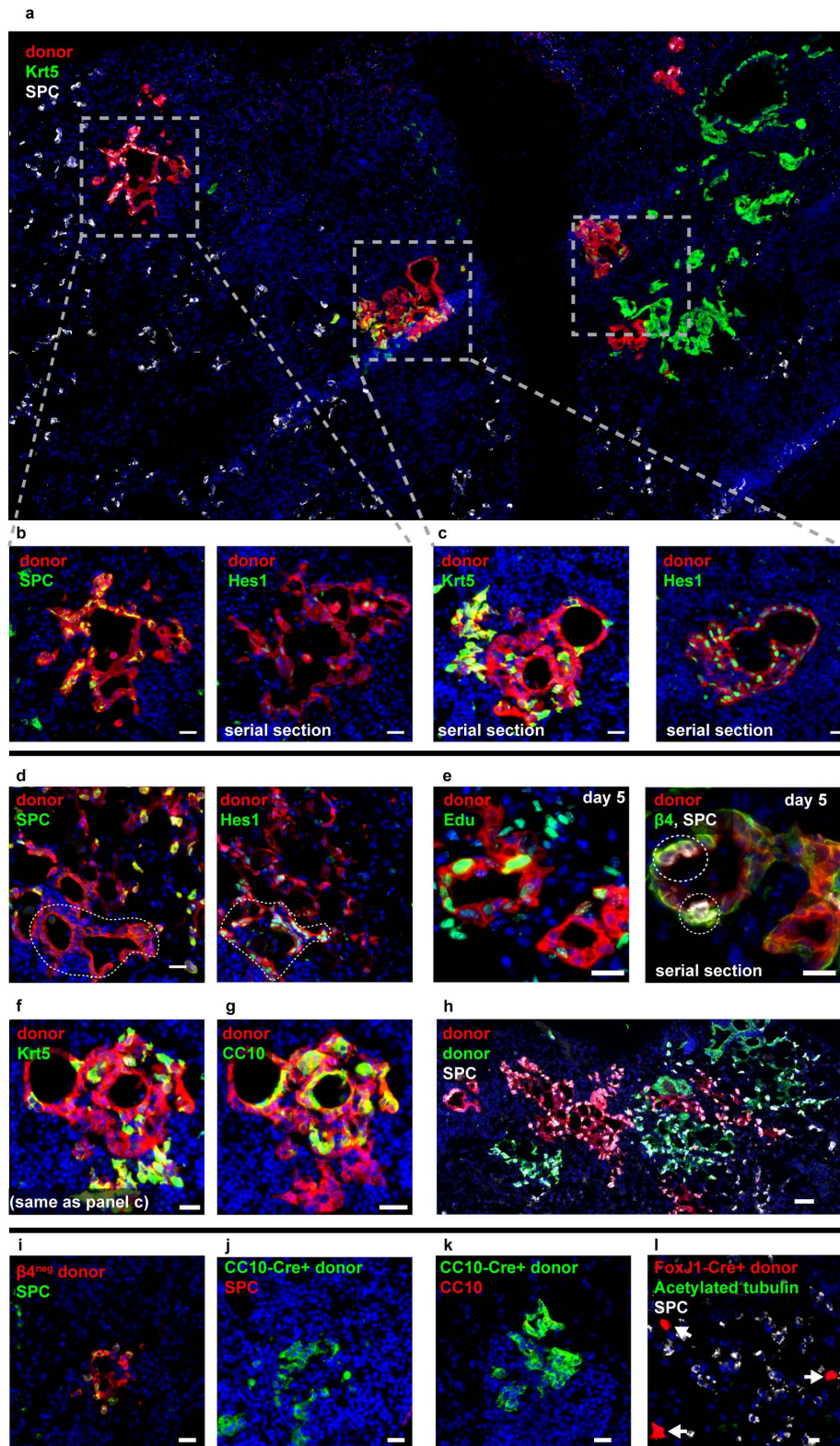
Extended Data Figure 4 | Krt5⁺ cells do not arise from CC10-expressing progenitors but rather upregulate CC10 during expansion. **a**, Krt5⁺ cells express detectable levels of CC10 (top) compared to isotype control (bottom) in alveolar clusters (**a**). **b**, Representative image of CC10-CreERT2 lineage trace in which waiting only 7 days after tamoxifen administration before influenza

injury results in significant labelling of Krt5⁺ cells (quantified in Fig. 1d). **c**, Strong CC10 expression in Krt5-CreERT2-traced (tdTomato⁺) cells by day 22 after influenza. For comparison, see single channel images (**c**, right and bottom) of the same region. Scale bars, 20 μm.



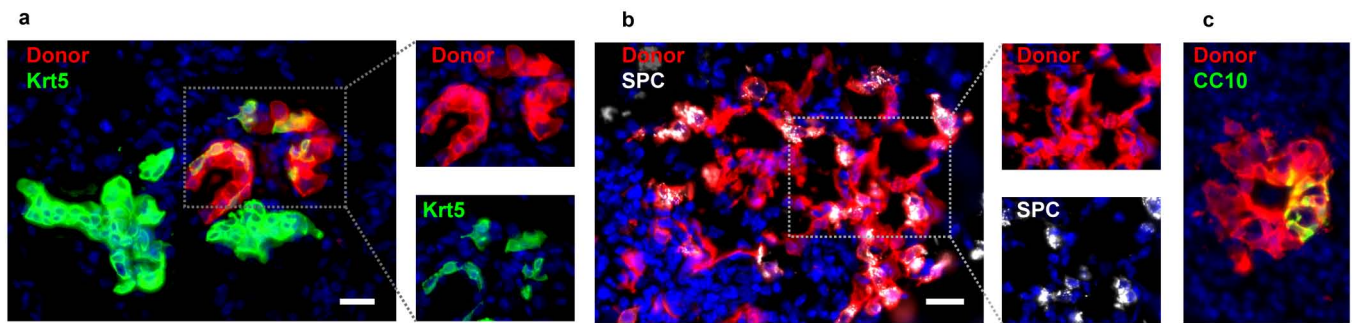
Extended Data Figure 5 | Heterogeneity of the LNEP-containing CC10[−]β₄⁺ population. **a**, Rare Krt5-CreERT2-traced (tdTomato⁺) cells were observed in uninjured distal lung airways that lacked Krt5 staining compared to trachea basal cells (inset) in the same section. All distal tdTomato⁺ cells express ΔNp63 but most ΔNp63⁺ cells are untraced (see Fig. 2c). **b**, Cytospins of sorted CC10[−]β₄⁺ cells reveal the presence of abundant multiciliated cells (green, acetylated tubulin⁺) and a small fraction of ΔNp63⁺ cells (red). **c**, Quantitative reverse transcriptase PCR (qRT-PCR) analysis of mature lineage genes and genes of interest in all populations. *n* = 3 biological replicates; data are

mean ± s.d. **d**, Principal component analysis plot of cells sequenced in Fig. 2b, demonstrating that p63⁺ cells in the CC10[−]β₄⁺ population (outlined, asterisk) cluster with multi-ciliated cells. **e**, CD200 is not expressed by Foxj1-CreERT2-labelled multi-ciliated cells, highlighting its use in excluding such cells. **f**, Cytospin of Foxj1-CreERT2-labelled β₄⁺ cells demonstrating reliable selection for multi-ciliated cells (198 cells quantified). **g**, Gating on CD14 expression within the EpcAM⁺β₄⁺CD200⁺ population excludes CC10-expressing club cells. Scale bars, 20 μm.

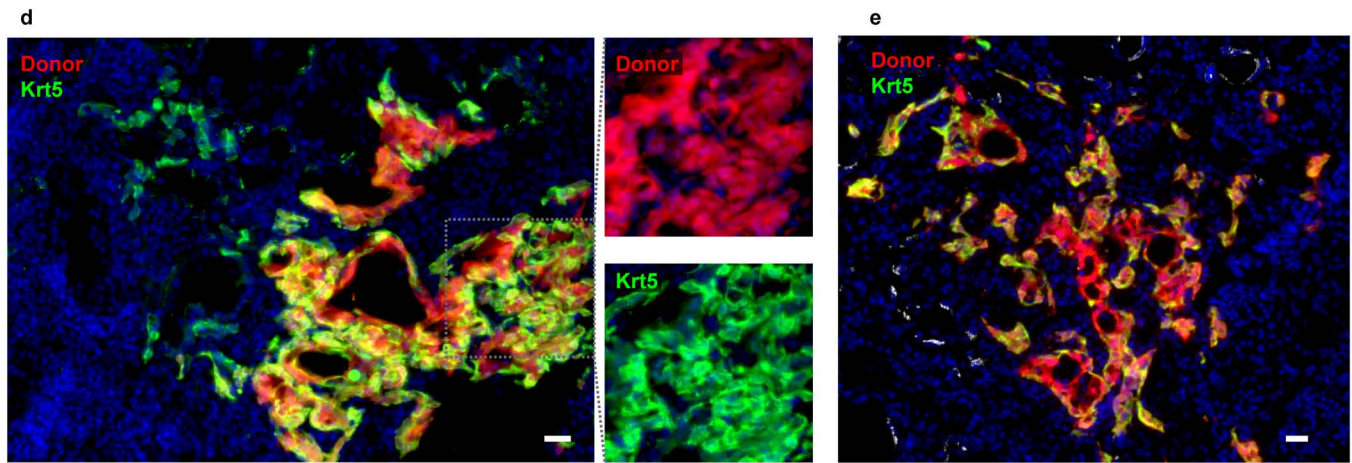


Extended Data Figure 6 | Orthotopic transplantation of LNEPs reveals their multipotency and differentiation appropriate to the local microenvironment. **a**, Several distinct areas of LNEP engraftment (red) reflect differentiation in response to location. Left dashed box demonstrates SPC expression in engrafted cells with nearby endogenous SPC-expressing cells (white); far right dashed box demonstrates Krt5 expression in engrafted cells and nearby endogenous Krt5-expressing cells (green). **b**, **c**, Cells in regions of SPC⁺ differentiation (**b**) lack Hes1 expression (right), whereas those in areas of Krt5⁺ differentiation (**c**) strongly express Hes1 (right). **d**, Distinct areas of LNEP engraftment demonstrate an inverse relationship between SPC

expression (left) and Hes1 expression (right) in probable single clones. **e**, Examination of transplanted cells 5 days after engraftment demonstrate abundant Edu incorporation (see Methods) indicative of proliferation. At this time point cells can be identified co-expressing β_4 and SPC (right, circled). **f**, **g**, Krt5⁺ cells and CC10⁺ cells were often found clustered in single regions of engraftment. **h**, Many engrafted cells in Fig. 2e are also SPC positive. **i**, β_4 ⁺ type II cells engraft in small clusters and only express SPC. **j**, **k**, CC10⁺ cells engraft but do not express SPC, CC10 or Krt5. **l**, Multi-ciliated cells engraft but only persist as isolated single cells, losing acetylated tubulin expression. Scale bars, 100 μ m (**a**) and 20 μ m (**b**–**l**).

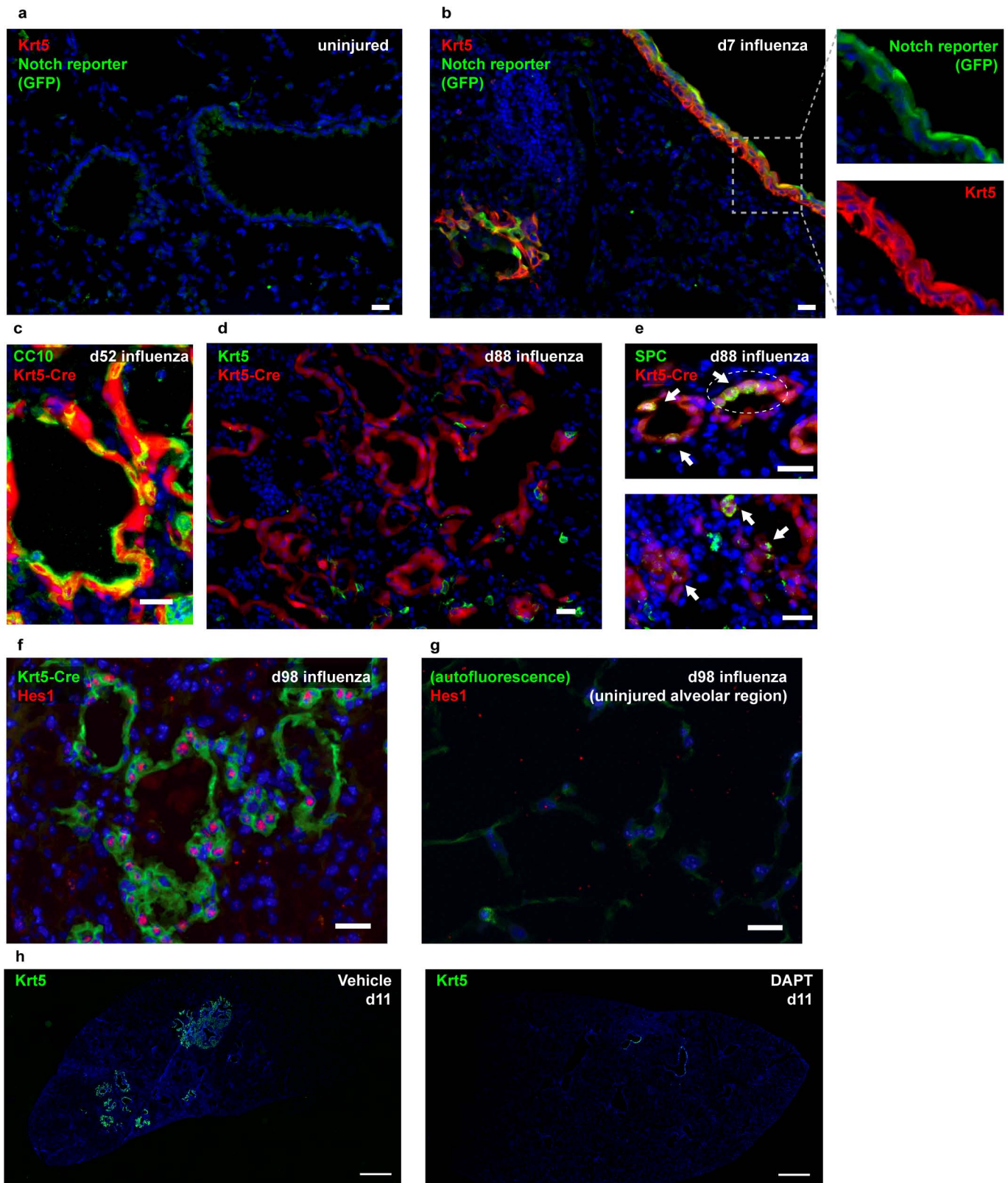
β_4^+ CD200⁺ CD14⁺ transplant

Krt5-CreERT2 / tdTomato transplant



Extended Data Figure 7 | Transplantation of β_4^+ CD14⁺ CD200⁺ and Krt5-CreERT2-traced cells recapitulates multipotency of the heterogeneous CC10⁻ β_4^+ population. **a**, Single channels images from Fig. 2h demonstrate Krt5 expression in transplanted β_4^+ CD14⁺ CD200⁺ cells. **b**, **c**, Transplanted β_4^+ CD14⁺ CD200⁺ can also differentiate towards type II

cells (**b**) and club cells (**c**). **d**, **e**, Transplantation of rare Krt5-CreERT2-traced cells from uninjured mice resulting in donor-derived Krt5⁺ cell expansion indistinguishable from endogenous expansion. Images in **d** and **e** are representative images from four attempted transplants, two of which exhibited engraftment in two or four individual lobes. Scale bars, 20 μ m.

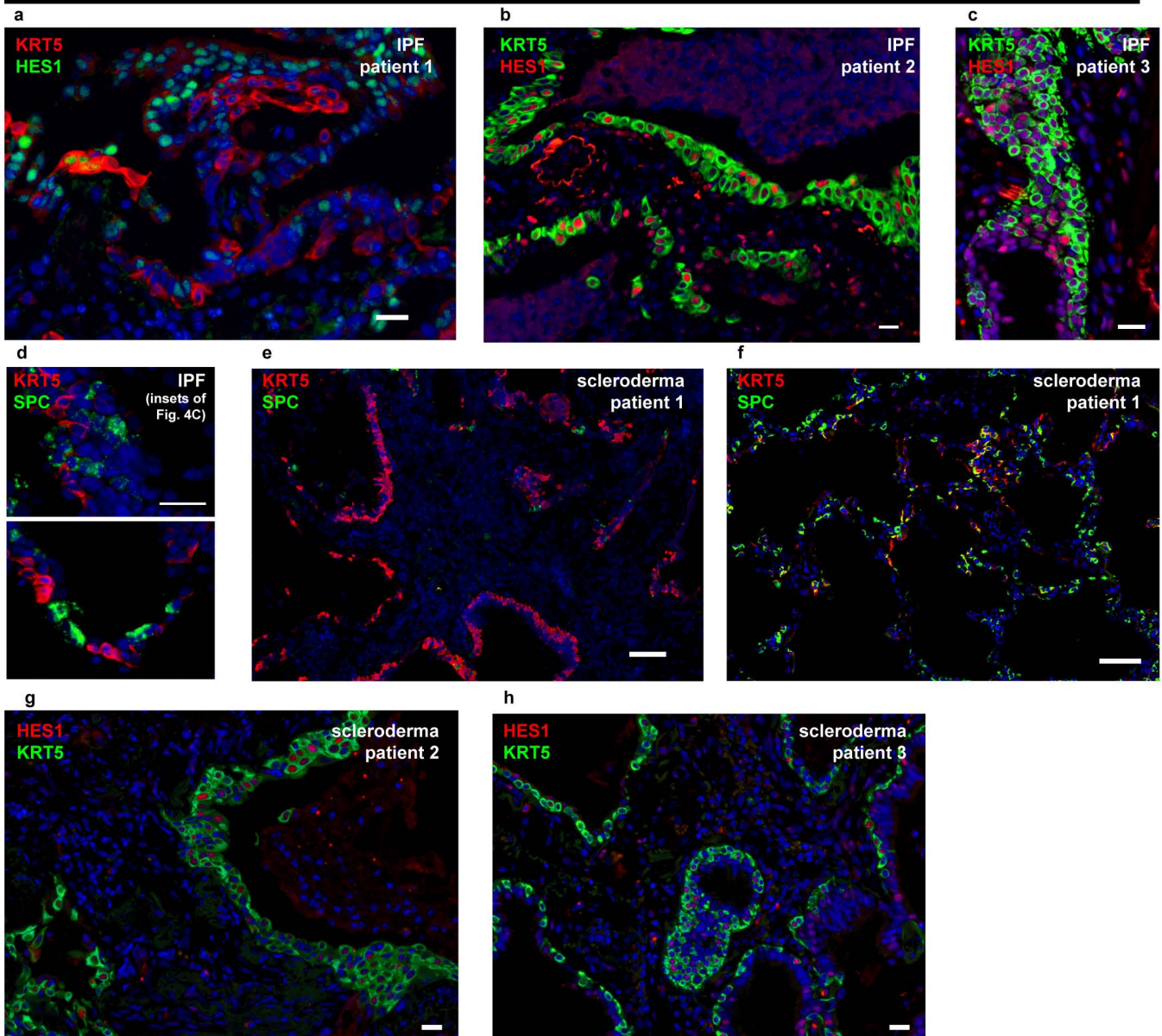


Extended Data Figure 8 | Notch activity in normal and injured lung.

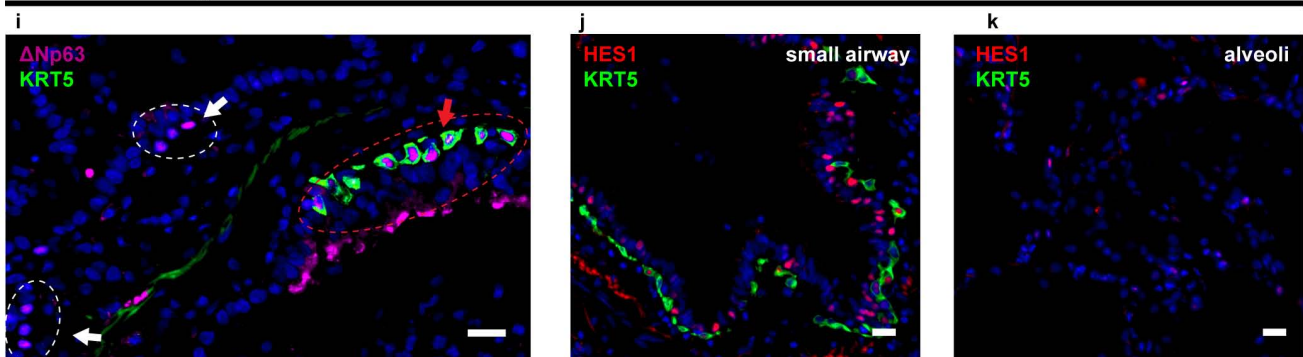
a, Uninjured Notch reporter mice (Cp-eGFP) show dim GFP in small airways and no detectable GFP in alveoli. **b**, Krt5⁺ cells arising in distal airways express GFP in Notch reporter mice 7 days after influenza infection. **c**, **d**, Some Krt5⁺ cells persist within Krt5-CreERT2-labelled (tdTomato⁺) cysts (**d**) long-term (day 88) after influenza injury, and many traced cells express CC10 (**c**).

e, Cysts rarely contain SPC⁺ type II cells (arrows). **f**, **g**, Hes1 expression is maintained in Krt5-CreERT2-traced (GFP⁺) cyst cells 98 days after influenza (**f**) but is absent in normal alveolar parenchyma from the same mice (**g**). **h**, Representative images of Krt5⁺ cell expansion in vehicle- (left) or DAPT- (right) treated mice at day 11 after influenza, quantified in Fig. 3g. Scale bars, 20 μ m (**a–g**) and 100 μ m (**h**).

diseased human lung

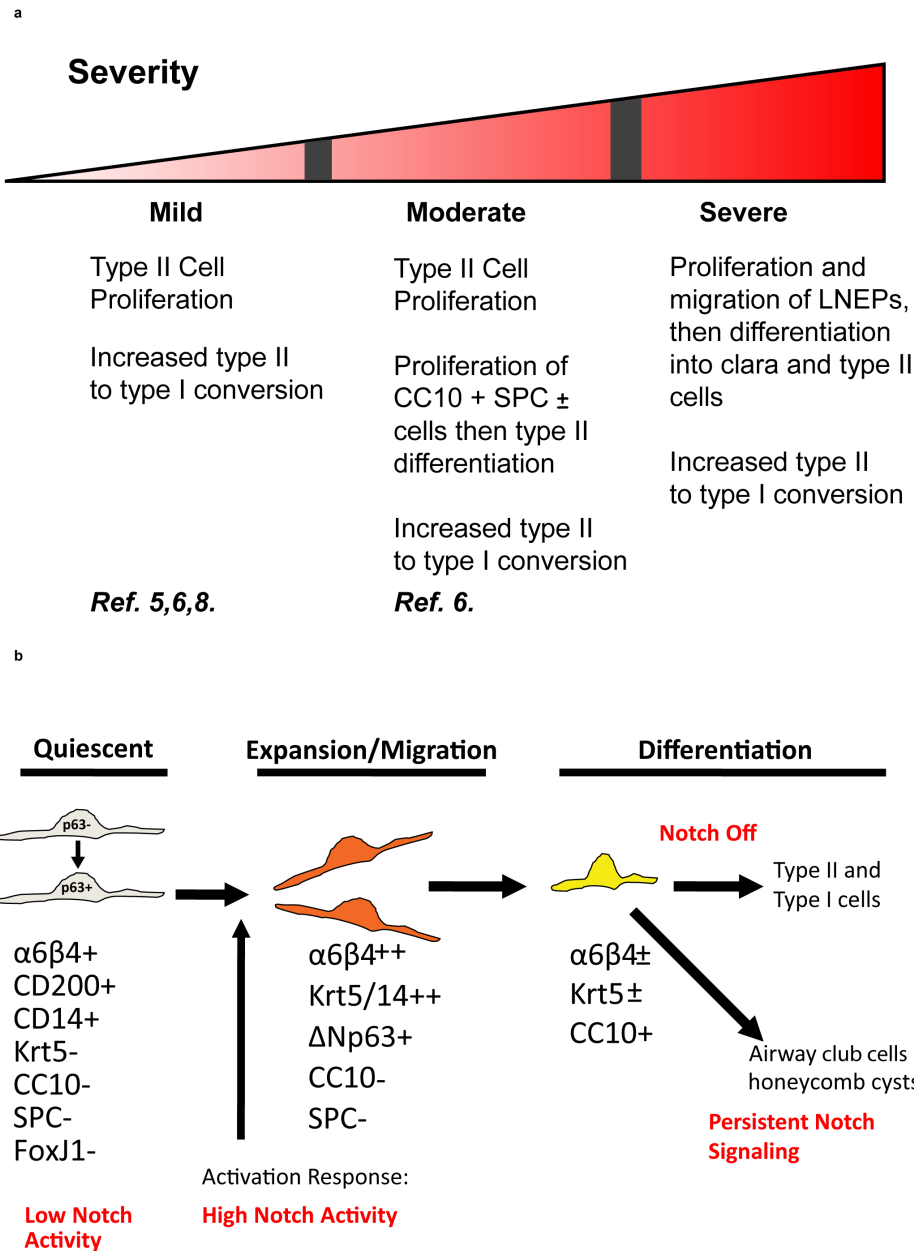


normal human lung



Extended Data Figure 9 | IPF and scleroderma lungs both contain HES1⁺ honeycomb cysts, but scleroderma lungs also possess SPC and KRT5 co-expressing cells. Normal human lungs contain putative LNEPs and lack HES1 in alveoli. **a–d**, Honeycomb cysts in several IPF lungs; many KRT5⁺ cells as well as surrounding cystic epithelium demonstrate strong nuclear HES1 signal. **e**, Region of scleroderma honeycombing similar to IPF lung. **f**, Scleroderma subpleural alveolar region with type II cell hyperplasia demonstrating cells

co-expressing SPC and KRT5. **g, h**, Cystic epithelium in scleroderma lungs expresses HES1 as in IPF. **i**, KRT5⁺ ΔNp63⁺ cells (white outlines) distinct from KRT5⁺ ΔNp63⁺ basal cells (red outlines) are present in distal airways. **j, k**, HES1 staining is apparent in small airways of normal lung (**j**) but very low in alveolar parenchyma (**k**). All images are from patient samples in addition to those shown in Fig. 4. Scale bars, 20 μm (**a–d, g–k**) and 100 μm (**e, f**).



Extended Data Figure 10 | Hierarchical cellular responses to injury severity and Notch-regulated LNEP dynamics. **a**, Distinct epithelial cell types contribute to regeneration depending on the severity of parenchymal injury. Examples of each are referenced. **b**, Notch signalling regulates the activation,

expansion and differentiation of LNEPs. Notch is required for activation and maintenance of LNEPs. Alveolar differentiation requires subsequent loss of Notch activity, whereas persistent Notch results in either airway differentiation or abnormal cystic honeycombing.

IAPP-driven metabolic reprogramming induces regression of *p53*-deficient tumours *in vivo*

Avinashnarayan Venkatanarayan^{1,2,3,4}, Payal Raulji^{1,2}, William Norton⁵, Deepavali Chakravarti^{1,2,3,4}, Cristian Coarfa⁶, Xiaohua Su^{1,2,4}, Santosh K. Sandur^{1,2,4,7}, Marc S. Ramirez⁸, Jaehuk Lee⁸, Charles V. Kingsley⁸, Eliot F. Sananikone^{1,2,3,4}, Kimal Rajapakse⁶, Katherine Naff⁵, Jan Parker-Thornburg⁹, James A. Bankson⁸, Kenneth Y. Tsai^{2,10}, Preethi H. Gunaratne¹¹ & Elsa R. Flores^{1,2,3,4}

TP53 is commonly altered in human cancer, and *TP53* reactivation suppresses tumours *in vivo* in mice^{1,2} (*TP53* and *TP53* are also known as *p53*). This strategy has proven difficult to implement therapeutically, and here we examine an alternative strategy by manipulating the *p53* family members, *TP63* and *TP73* (also known as *p63* and *p73*, respectively). The acidic transactivation-domain-bearing (TA) isoforms of *p63* and *p73* structurally and functionally resemble *p53*, whereas the ΔN isoforms (lacking the acidic transactivation domain) of *p63* and *p73* are frequently overexpressed in cancer and act primarily in a dominant-negative fashion against *p53*, TAp63 and TAp73 to inhibit their tumour-suppressive functions^{3–8}. The *p53* family interacts extensively in cellular processes that promote tumour suppression, such as apoptosis and autophagy^{9–14}, thus a clear understanding of this interplay in cancer is needed to treat tumours with alterations in the *p53* pathway. Here we show that deletion of the ΔN isoforms of *p63* or *p73* leads to metabolic reprogramming and regression of *p53*-deficient tumours through upregulation of *IAPP*, the gene that encodes amylin, a 37-amino-acid peptide co-secreted with insulin by the β cells of the pancreas. We found that *IAPP* is causally involved in this tumour regression and that amylin functions through the calcitonin receptor (CalcR) and receptor activity modifying protein 3 (RAMP3) to inhibit glycolysis and induce reactive oxygen species and apoptosis. Pramlintide, a synthetic analogue of amylin that is currently used to treat type 1 and type 2 diabetes, caused rapid tumour regression in *p53*-deficient thymic lymphomas, representing a novel strategy to target *p53*-deficient cancers.

Using $\Delta Np63$ (ref. 15) and $\Delta Np73$ conditional knockout mice (Extended Data Fig. 1a, b), we generated $\Delta Np63^{+/-}$ and $\Delta Np73^{+/-}$ mice (Extended Data Fig. 1c–f). To ask whether the ΔN isoforms of *p63* and *p73* act as oncogenes *in vivo* by interacting with *p53*, $\Delta Np63^{+/-};p53^{-/-}$ and $\Delta Np73^{+/-};p53^{-/-}$ mice were aged for the development of thymic lymphomas, which form in nearly all *p53*^{-/-} mice¹⁶. We found a remarkable diminution in the number and size of thymic lymphomas in $\Delta Np63^{+/-};p53^{-/-}$ and $\Delta Np73^{+/-};p53^{-/-}$ mice, leading to an extended lifespan (Extended Data Fig. 2a–c) and suggesting that the ΔN isoforms of *p63* and *p73* restrain a tumour suppressive program that can compensate for *p53* function.

We found that TAp63 and TAp73 were upregulated in thymic lymphomas from $\Delta Np63^{+/-};p53^{-/-}$ and $\Delta Np73^{+/-};p53^{-/-}$ mice (Extended Data Fig. 2d, e) along with an upregulation of apoptosis (Extended Data Fig. 2f–j) and senescence (Extended Data Fig. 2k–o). We also examined thymocytes from 4-week-old mice after treatment with 10 Gy gamma irradiation, a dose that is known to elicit *p53*-dependent apoptosis^{9,17}.

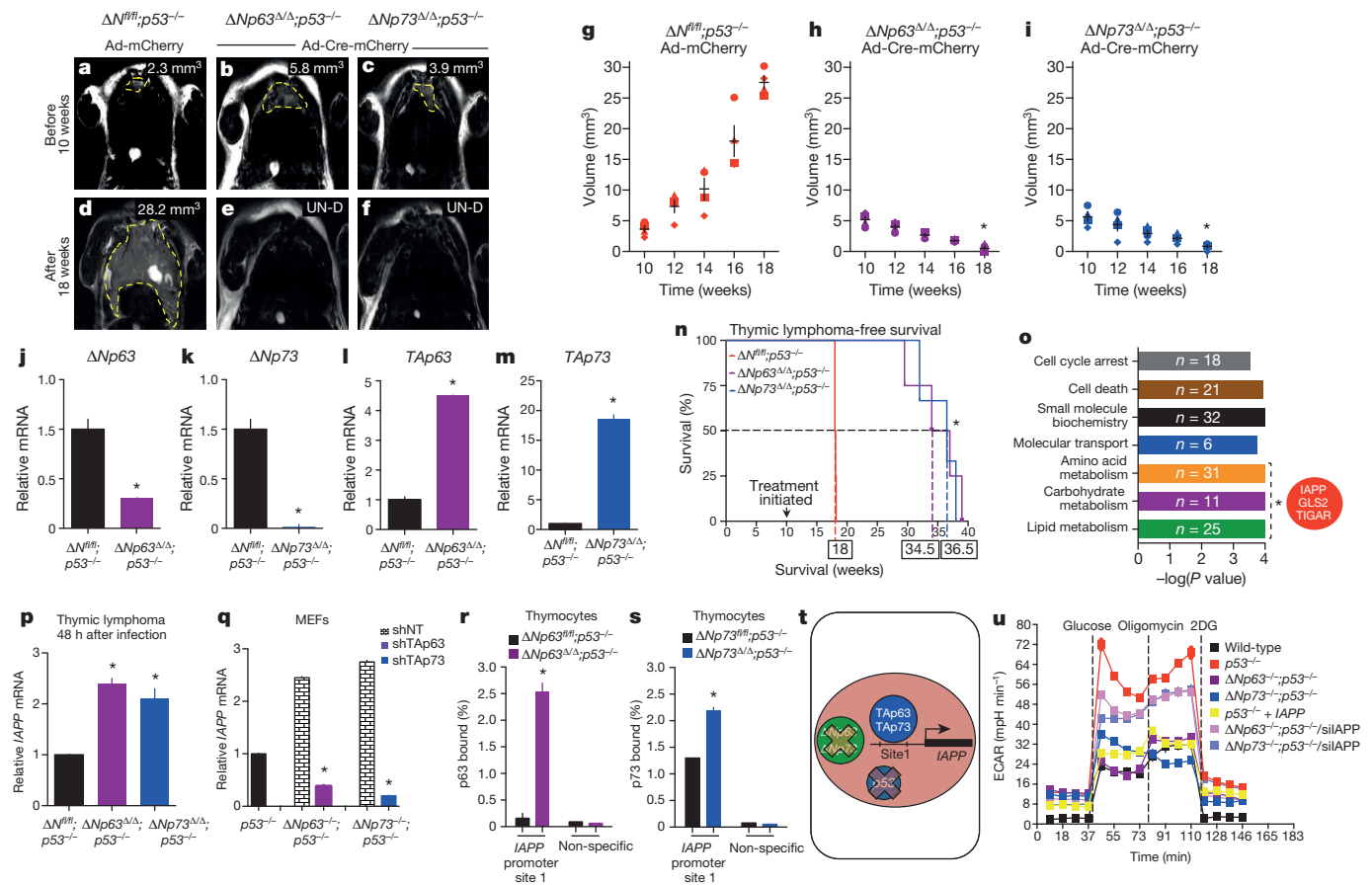
Indeed, TAp63 and TAp73 are higher in $\Delta Np63^{+/-};p53^{-/-}$ and $\Delta Np73^{+/-};p53^{-/-}$ thymocytes, which was further exacerbated after gamma irradiation (Extended Data Fig. 3a–c) with an increase in apoptosis (Extended Data Fig. 3d–h) and senescence (Extended Data Fig. 3i–m).

To determine whether TAp63 or TAp73 compensate for *p53* function in tumours *in vivo*, we acutely removed $\Delta Np63$ or $\Delta Np73$ by intratumoral infection with adenovirus-Cre-mCherry (Extended Data Fig. 4a–d and Fig. 1a–f) in $\Delta Np63^{fl/fl};p53^{-/-}$ and $\Delta Np73^{fl/fl};p53^{-/-}$ at 10 weeks of age. Tumours were 2.3–5.8 mm³ in size at the time of infection and monitored weekly by magnetic resonance imaging (MRI; Fig. 1a–i). Mice deficient for either $\Delta Np63$ or $\Delta Np73$ and *p53* showed marked decreases in tumour burden (Fig. 1h, i). The reduction of $\Delta Np63$ and $\Delta Np73$ expression resulted in increased expression of TAp63 and TAp73 (Fig. 1j–m and Extended Data 4d) and increased apoptosis (Extended Data Fig. 4e–h) and senescence (Extended Data Fig. 4i–k). $\Delta Np63^{fl/fl};p53^{-/-}$ and $\Delta Np73^{fl/fl};p53^{-/-}$ mice also had an increased lifespan (Fig. 1n). We found differences in CD4/CD8-positive cells in young mice (4 weeks) (Extended Data Fig. 4l–p), indicating that changes in T-cell development may lead to a lower tumour incidence in double-mutant mice. Indeed, we found that *p53*^{-/-} thymic lymphomas are composed primarily of CD4/CD8 double-positive thymocytes while the $\Delta Np63^{fl/fl};p53^{-/-}$ and $\Delta Np73^{fl/fl};p53^{-/-}$ lymphomas contain very few CD4/CD8 double-positive thymocytes (Extended Data Fig. 4q–t). Lastly, we asked whether thymic stromal cells contribute to the apoptosis in the regressing lymphomas. We sorted CD45-positive cells to select for T lymphocytes in *p53*^{-/-}, $\Delta Np63^{fl/fl};p53^{-/-}$ and $\Delta Np73^{fl/fl};p53^{-/-}$ mice and infected them with adenovirus-Cre (Extended Data Fig. 4u). $\Delta Np63^{fl/fl};p53^{-/-}$ and $\Delta Np73^{fl/fl};p53^{-/-}$ thymocytes underwent apoptosis independent of the presence of the stromal cells (Extended Data Fig. 4v). These data indicate that inhibition of the ΔN isoforms of *p63* and *p73* serves to upregulate TAp63 and TAp73 to compensate for loss of *p53* in tumour suppression.

We found that the ΔN isoforms of *p63* and *p73* bind to the promoters of the TA isoforms of *p63* and *p73*, suggesting that the ΔN isoforms of *p63* and *p73* can transcriptionally repress TAp63 and TAp73 transcription (Extended Data Fig. 5a–i). We also found that the increase in apoptosis and cellular senescence was dependent on TAp63 and TAp73 (Extended Data Fig. 5j–q).

We performed RNA sequencing of lymphomas after infection with Ad-mCherry ($\Delta Np63^{fl/fl};p53^{-/-}$ and $\Delta Np73^{fl/fl};p53^{-/-}$) and Ad-Cre-mCherry ($\Delta Np63^{fl/fl};p53^{-/-}$ and $\Delta Np73^{fl/fl};p53^{-/-}$) and found that thymic lymphomas from mice deficient for *p53* and $\Delta Np63$ clustered with those from mice deficient for *p53* and $\Delta Np73$ (Extended Data Fig. 6a). Ingenuity pathway analysis (IPA) (Fig. 1q) revealed genes involved in

¹Department of Molecular and Cellular Oncology, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas 77030, USA. ²Department of Translational Molecular Pathology, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas 77030, USA. ³Graduate School of Biomedical Sciences, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas 77030, USA. ⁴Metastasis Research Center, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas 77030, USA. ⁵Department of Veterinary Medicine and Surgery, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas 77030, USA. ⁶Department of Molecular and Cellular Biology, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas 77030, USA. ⁷Radiation Biology & Health Sciences Division, Bhabha Atomic Research Center, Mumbai 400085, India. ⁸Department of Imaging Physics, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas 77030, USA. ⁹Department of Genetics, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas 77030, USA. ¹⁰Department of Dermatology, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas 77030, USA. ¹¹Department of Biology and Biochemistry, University of Houston, Houston, Texas 77204, USA.



metabolism including TP53-inducible glycolysis and apoptosis regulator (*TIGAR*)¹⁸, and glutaminase 2 (*GLS2*)^{19,20}. While we found that *TIGAR* and *GLS2* were upregulated in either *ANP63*^{Δ/Δ}; *p53*^{−/−} or *ANP73*^{Δ/Δ}; *p53*^{−/−} thymic lymphomas, we identified a novel gene, islet amyloid polypeptide (*IAPP*) or amylin, which was upregulated by over fivefold in both double-mutant thymic lymphomas. *IAPP* limits glucose uptake, resulting in increased intracellular glucose-6-phosphate (G-6-P)²¹ levels and decreased glycolysis²¹. We validated *IAPP*, *TIGAR* and *GLS2* expression in thymic lymphomas derived from *ANP63*^{Δ/Δ}; *p53*^{−/−} and *ANP73*^{Δ/Δ}; *p53*^{−/−} mice and found that *IAPP* is expressed at levels over twofold higher in double-mutant mice (Fig. 1p and Extended Data Fig. 6b–d). *IAPP* and *GLS2* expression depend on *Tap63* and *Tap73* (Fig. 1q and Extended Data Fig. 6d). To determine whether *Tap63* or *Tap73* transcriptionally regulate *IAPP*, we performed chromatin immunoprecipitation in mouse embryonic fibroblasts (MEFs; Extended Data Fig. 6e–g) and thymocytes (Fig. 1r, s). We found that *Tap63* and *Tap73* bind to sites located in the promoter (site 1), 1,756 nucleotides upstream of the transcriptional start site, and intron 2 (site 2) of *IAPP*, 706 nucleotides downstream of the transcriptional start site (Extended Data Fig. 6e–g). Because a greater binding affinity of *Tap63* and *Tap73* was detected in the promoter region (site 1) of *IAPP*, we cloned this site into a luciferase reporter gene and also mutated this site (Extended Data Fig. 6h–k). Only the luciferase reporter gene containing wild-type *IAPP* promoter site 1 was transactivated by *Tap63* and *Tap73* whereas the

mutant version was not. Taken together, these data indicate that *IAPP* is a transcriptional target gene of *Tap63* and *Tap73* (Fig. 1t). Expression of *IAPP* in *p53*^{−/−} MEFs resulted in low levels of glycolysis comparable to that in *ANP63*^{−/−}; *p53*^{−/−} and *ANP73*^{−/−}; *p53*^{−/−} MEFs (Extended Data Fig. 6l–m and Fig. 1u). Conversely, when we knocked down *IAPP* in *ANP63*^{−/−}; *p53*^{−/−} and *ANP73*^{−/−}; *p53*^{−/−} MEFs, the levels of glycolysis were similar to that of *p53*^{−/−} MEFs (Fig. 1u) indicating that *IAPP* inhibits glycolysis. *In vivo*, we detected massive tumour regression in *ANP63*^{fl/fl}; *p53*^{−/−} or *ANP73*^{fl/fl}; *p53*^{−/−} thymic lymphomas treated with *IAPP* (Extended Data Fig. 7a and Fig. 2a, b, h, i, o, p), $P < 0.05$. Conversely, in *ANP63*^{Δ/Δ}; *p53*^{−/−} and *ANP73*^{Δ/Δ}; *p53*^{−/−} thymic lymphomas treated with Ad-shIAPP-mCherry the tumours continued to grow comparable to that of *p53*^{−/−} thymic lymphomas (Fig. 2a–k, o–r), $P > 0.05$ at 13 weeks. Additionally, *p53*^{−/−} mice treated with Ad-IAPP had an extended tumour-free survival period compared to *p53*^{−/−} mice or *ANP63*^{Δ/Δ}; *p53*^{−/−} and *ANP73*^{Δ/Δ}; *p53*^{−/−} mice treated intratumorally with Ad-shIAPP-mCherry (Extended Data Fig. 7a, b), indicating that *IAPP* is a tumour suppressor gene and is causally involved in the *in vivo* effects seen upon inactivation of *ANP63* or *ANP73*. Given that pramlintide, a synthetic analogue of amylin, is used to treat type I and type II diabetes²², we treated thymic lymphomas in *ANP63*^{fl/fl}; *p53*^{−/−} and *ANP73*^{fl/fl}; *p53*^{−/−} mice. Indeed, three-weekly intratumoral injections resulted in rapid tumour regression (Fig. 2e, l, s), $P < 0.005$ at 13 weeks. This effect was exacerbated by systemic intravenous treatment with

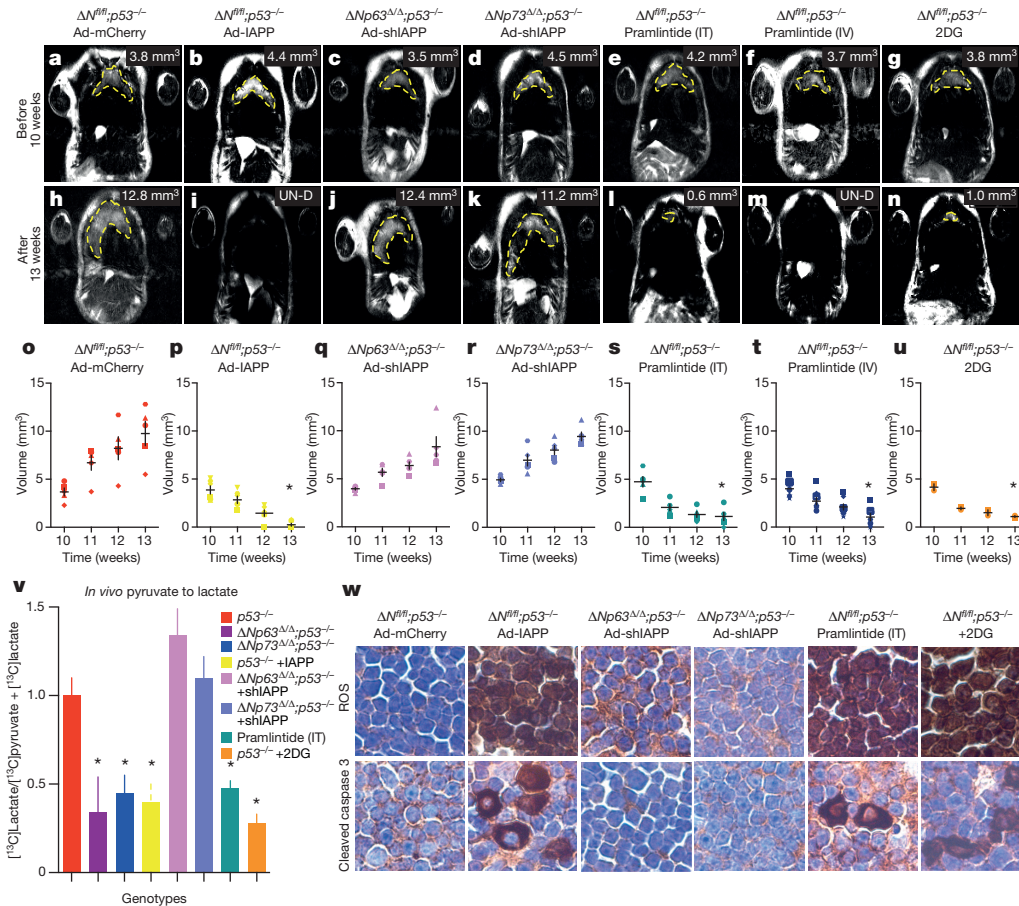


Figure 2 | IAPP is causally involved in tumorigenesis suppression in p53-deficient thymic lymphomas. a–n, Thymic lymphomas were infected with adenovirus (Ad)-mCherry (a, h), Ad-IAPP-mCherry (+IAPP) (b, i), Ad-shIAPP-mCherry (c, d, j, k), or treated with pramlintide intratumorally (IT) (e, l) or intravenously (IV) (f, m), or with 2DG (g, n). Yellow dashed lines indicate tumour. Volume of tumour shown. UN-D, undetectable.

pramlintide (Fig. 2f, m, t and Extended Data Fig. 7c–q), $P < 0.005$, similar to that seen in tumours treated with a known inhibitor of glycolysis, 2-deoxy-D-glucose (2DG; Fig. 2g, n, u). These data provide preclinical *in vivo* evidence that pramlintide can be used to effectively treat p53-deficient tumours. Using *in vivo* dynamic magnetic resonance spectroscopy to measure the conversion of hyperpolarized [1-¹³C]pyruvate to lactate as a proxy of glycolysis within the tumours, we found a marked reduction in glycolysis in ΔNp63/p53 and ΔNp73/p53 double-deficient mice and after introducing IAPP into p53^{-/-} thymic lymphomas similar to tumours treated with 2DG (Fig. 2v). ΔNp63^{Δ/Δ};p53^{-/-} and ΔNp73^{Δ/Δ};p53^{-/-} thymic lymphomas infected with a short hairpin RNA for IAPP exhibited levels of glycolysis similar to those found in p53^{-/-} thymic lymphomas (Fig. 2v). Pramlintide also inhibits glycolysis in tumours (Fig. 2v).

IAPP has been shown to induce reactive oxygen species (ROS) and activate apoptosis^{23,24}. We found a marked increase in the levels of ROS and apoptosis in thymic lymphomas expressing IAPP or treated with pramlintide or 2DG, whereas neither ROS nor apoptosis occurred upon inactivation of IAPP in thymic lymphomas from ΔNp63^{Δ/Δ};p53^{-/-} and ΔNp73^{Δ/Δ};p53^{-/-} mice (Fig. 2w), indicating that upregulation of IAPP inhibits glycolysis similarly to 2DG and leads to oxidative stress that triggers apoptosis. While high levels of ROS are not commonly triggered by inhibition of glycolysis, nutrient deprivation or excess can result in the accumulation of ROS. Additionally, cancer cells tightly regulate ROS by acquiring additional mutations and compensatory mechanisms often ensue and may be at play in the thymic lymphoma cells that acutely downregulate glycolysis by IAPP²⁵.

o–u, Quantification of the indicated thymic lymphomas, $n = 5$ mice per group. Significance indicated by the asterisks, $P < 0.005$. v, Quantification of *in vivo* pyruvate to lactate conversion using dynamic magnetic resonance spectroscopy as a measurement of glycolysis, $n = 3$ mice, $P < 0.005$. w, Immunohistochemistry for reactive oxygen species (ROS) or cleaved caspase 3. Positive nuclei are brown.

To extend our findings to human cancer where p53 is altered in the majority of cases, we analysed human cancer cell lines containing p53 deletions or mutations. We used short interfering RNA (siRNA) to knockdown ΔNp63 or ΔNp73 in cells derived from a lung adenocarcinoma (H1299) (Fig. 3a). Downregulation of ΔNp63 or ΔNp73 resulted in upregulation of TAp63, TAp73 and IAPP (Fig. 3a) and an increase in apoptosis and decrease in cell proliferation (Fig. 3b and Extended Data Fig. 8a–d). To ask whether IAPP can also inhibit glycolysis in human cancer cell lines, we transfected H1299 cells with siΔNp63, siΔNp73 or IAPP (Fig. 3a). Knockdown of ΔNp63 or ΔNp73 or expression of IAPP resulted in an inhibition of glycolysis (Fig. 3c, d) and glucose uptake (Extended Data Fig. 8e, g), accumulation of ROS (Fig. 3d–f), and induction of apoptosis (Fig. 3d, g, h). We inhibited ROS in these cells using N-acetyl-L-cysteine (NAC) and observed no apoptosis (Fig. 3d–h). Previous studies have indicated that IAPP inhibits glycolysis by increasing intracellular G-6-P in turn leading to an inhibition of hexokinase^{21,26}. We measured the levels of intracellular G-6-P in H1299 cells and found that cells expressing high levels of IAPP (H1299-siΔNp63, H1299-siΔNp73, or H1299+IAPP) also had high levels of G-6-P while knockdown of IAPP resulted in a diminution in G-6-P (Extended Data Fig. 8f, g). Overexpression of glucose hexokinase II (HKII) led to a rescue of the glycolytic capacity of H1299 cells expressing siΔNp63 or siΔNp73 to levels similar to those in parental H1299 cells (Fig. 3c–g). These results indicate that IAPP inhibits glycolysis through the inhibition of HKII. We found that treatment of H1299 cells with pramlintide led to similar effects on glycolysis and apoptosis (Fig. 3g–n). Taken together, these

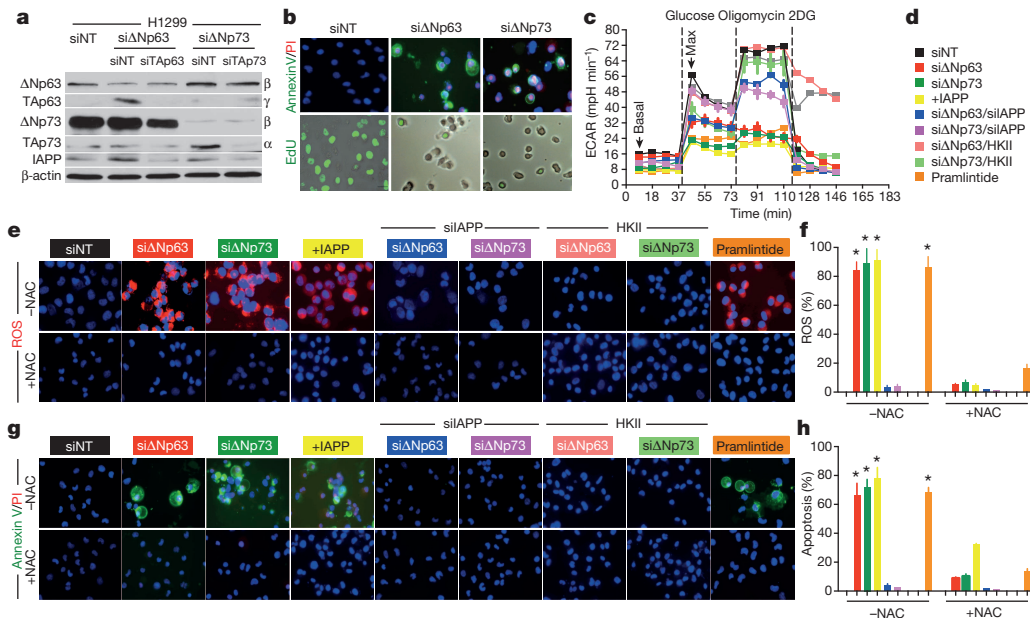


Figure 3 | IAPP inhibits glycolysis and induces ROS and apoptosis in p53-deficient human cancer cell lines. **a**, Representative western blot analysis, $n = 4$. **b**, Immunofluorescence for apoptosis and 5'-ethynyl-2'-deoxyuridine (EdU) incorporation. **c**, Extracellular acidification rate as a measure of

glycolysis, $n = 3$, $P < 0.005$. Legend in **d** is colour-coded and corresponds to panels **c**, **e–h**. **e–h**, Immunofluorescence and quantification for ROS (red) (**e**, **f**) or apoptosis (green or green/red) (**g**, **h**).

data demonstrate that IAPP and pramlintide inhibit glycolysis through the inhibition of HKII.

IAPP is a secreted protein and binds to the calcitonin receptor (CALCR) and RAMP3 (ref. 27). To determine whether IAPP functions through these receptors to inhibit glycolysis, secreted media from H1299 cells expressing siΔNp63 (siΔNp63^M) or siΔNp73 (siΔNp73^M), which contains secreted IAPP (Fig. 4a and Extended Data Fig. 9a, b), was added to H1299 cells resulting in inhibition of glycolysis (Fig. 4b) and induction of ROS and apoptosis (Fig. 4c, d). In contrast, when these media were used to treat H1299 cells with knockdown of CALCR or RAMP3, glycolysis was not inhibited and ROS and apoptosis were not induced (Fig. 4b–d), indicating that the CALCR and RAMP3 receptors are critical for IAPP function. We also treated the H1299 cells with media from H1299 cells expressing siΔNp63 (siΔNp63^M) or siΔNp73 (siΔNp73^M) and an amylin inhibitor, which led to high levels of glycolysis (Extended Data Fig. 9c) and low levels of ROS and apoptosis (Fig. 4c, d). IAPP causes activation of the NLRP3 inflammasome²⁸, which has been shown to be anti-tumorigenic in certain cancers via IL-18 processing²⁹. We blocked caspase-1 using an inhibitor and found that it prevented apoptosis of H1299 cells (Fig. 4d), demonstrating that pyroptosis may also be an important mechanism of action of IAPP.

To demonstrate the importance of the calcitonin receptor *in vivo*, we treated p53^{-/-} mice with thymic lymphomas at 10 weeks of age with pramlintide and a calcitonin receptor inhibitor (Fig. 4e–m) and found that this inhibition rendered pramlintide ineffective, demonstrating the importance of the calcitonin receptor for IAPP/amylin/pramlintide function (Fig. 4n). To further determine the anti-tumorigenic efficacy of

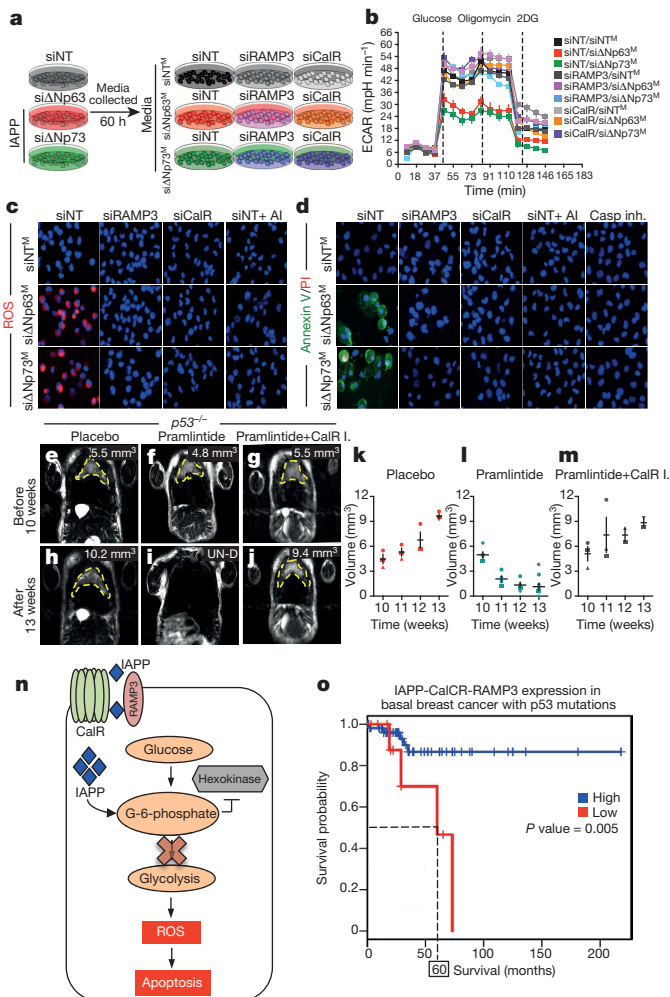


Figure 4 | Calcitonin and RAMP3 receptors are required for secreted IAPP to suppress tumorigenesis. **a**, Cartoon depicting treatment of cells expressing the indicated siRNAs and treated with media from the cells secreting IAPP on the left. **b**, Extracellular acidification rate (ECAR) in H1299 cells. **c**, **d**, Immunofluorescence for ROS (**c**) and apoptosis (**d**). **e–m**, MRI and quantification of thymic lymphomas treated with placebo (**e**, **h**, **k**), pramlintide (**f**, **i**, **l**), or pramlintide plus calcitonin inhibitor (CalR I.) (**g**, **j**, **m**), $n = 5$ mice. **n**, Cartoon of IAPP signalling through RAMP3 and calcitonin receptor (CALCR) to inhibit glycolysis and induce ROS and apoptosis. **o**, Kaplan–Meier curves from patients with p53 mutant tumours and co-expression of IAPP, RAMP3 and CALCR. Boxed numbers represent median survival.

pramlintide in cells with p53 deletions or mutations, we treated additional human cancer cell lines³⁰ with pramlintide and a calcitonin receptor inhibitor, resulting in increased glycolysis, decreased ROS and apoptosis (Extended Data Fig. 9d–i). We assessed patient survival using data from the Cancer Genome Atlas (TCGA) of patients with p53 mutations and found that co-expression of IAPP, CALCR and RAMP3 correlated with better patient survival in basal breast cancer (Fig. 4o), colorectal cancer and lung squamous cell carcinoma (Extended Data Fig. 9j, k).

Reactivation of p53 activity in tumours results in tumour suppression^{1,2}. We have focused on interactions between the three p53 family members and have revealed a novel strategy to target p53-deficient and mutant cancers through amylin-based therapies like pramlintide.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 3 December 2013; accepted 30 September 2014.

Published online 17 November 2014.

- Ventura, A. *et al.* Restoration of p53 function leads to tumour regression *in vivo*. *Nature* **445**, 661–665 (2007).
- Wang, Y. *et al.* Restoring expression of wild-type p53 suppresses tumor growth but does not cause tumor regression in mice with a p53 missense mutation. *J. Clin. Invest.* **121**, 893–904 (2011).
- Flores, E. R. *et al.* Tumor predisposition in mice mutant for p63 and p73: evidence for broader tumor suppressor functions for the p53 family. *Cancer Cell* **7**, 363–373 (2005).
- Su, X. *et al.* Tap63 suppresses metastasis through coordinate regulation of Dicer and miRNAs. *Nature* **467**, 986–990 (2010).
- Su, X., Chakravarti, D. & Flores, E. R. p63 steps into the limelight: crucial roles in the suppression of tumorigenesis and metastasis. *Nature Rev. Cancer* **13**, 136–143 (2013).
- Tomasini, R. *et al.* Tap73 knockout shows genomic instability with infertility and tumor suppressor functions. *Genes Dev.* **22**, 2677–2691 (2008).
- Yang, A. *et al.* p63, a p53 homolog at 3q27–29, encodes multiple products with transactivating, death-inducing, and dominant-negative activities. *Mol. Cell* **2**, 305–316 (1998).
- Su, X. *et al.* Tap63 is a master transcriptional regulator of lipid and glucose metabolism. *Cell Metab.* **16**, 511–525 (2012).
- Flores, E. R. *et al.* p63 and p73 are required for p53-dependent apoptosis in response to DNA damage. *Nature* **416**, 560–564 (2002).
- Kenzelmann Broz, D. *et al.* Global genomic profiling reveals an extensive p53-regulated autophagy program contributing to key p53 responses. *Genes Dev.* **27**, 1016–1031 (2013).
- Di Como, C. J., Gaididon, C. & Prives, C. p73 function is inhibited by tumor-derived p53 mutants in mammalian cells. *Mol. Cell. Biol.* **19**, 1438–1449 (1999).
- Gaididon, C., Lokshin, M., Ahn, J., Zhang, T. & Prives, C. A subset of tumor-derived mutant forms of p53 down-regulate p63 and p73 through a direct interaction with the p53 core domain. *Mol. Cell. Biol.* **21**, 1874–1887 (2001).
- Lang, G. A. *et al.* Gain of function of a p53 hot spot mutation in a mouse model of Li-Fraumeni syndrome. *Cell* **119**, 861–872 (2004).
- Olive, K. P. *et al.* Mutant p53 gain of function in two mouse models of Li-Fraumeni syndrome. *Cell* **119**, 847–860 (2004).
- Chakravarti, D. *et al.* Induced multipotency in adult keratinocytes through down-regulation of *ΔNp63* or *DGCR8*. *Proc. Natl Acad. Sci. USA* **111**, E572–E581 (2014).
- Jacks, T. *et al.* Tumor spectrum analysis in p53-mutant mice. *Curr. Biol.* **4**, 1–7 (1994).
- Attardi, L. D., de Vries, A. & Jacks, T. Activation of the p53-dependent G1 checkpoint response in mouse embryo fibroblasts depends on the specific DNA damage inducer. *Oncogene* **23**, 973–980 (2004).
- Bensaad, K. *et al.* TIGAR, a p53-inducible regulator of glycolysis and apoptosis. *Cell* **126**, 107–120 (2006).
- Li, T. *et al.* Tumor suppression in the absence of p53-mediated cell-cycle arrest, apoptosis, and senescence. *Cell* **149**, 1269–1283 (2012).
- Suzuki, S. *et al.* Phosphate-activated glutaminase (GLS2), a p53-inducible regulator of glutamine metabolism and reactive oxygen species. *Proc. Natl Acad. Sci. USA* **107**, 7461–7466 (2010).
- Castle, A. L., Kuo, C. H., Han, D. H. & Ivy, J. L. Amylin-mediated inhibition of insulin-stimulated glucose transport in skeletal muscle. *Am. J. Physiol.* **275**, E531–E536 (1998).
- Edelman, S., Maier, H. & Wilhelm, K. Pramlintide in the treatment of diabetes mellitus. *BioDrugs* **22**, 375–386 (2008).
- Mattson, M. P. & Goodman, Y. Different amyloidogenic peptides share a similar mechanism of neurotoxicity involving reactive oxygen species and calcium. *Brain Res.* **676**, 219–224 (1995).
- Schubert, D. *et al.* Amyloid peptides are toxic via a common oxidative mechanism. *Proc. Natl Acad. Sci. USA* **92**, 1989–1993 (1995).
- Cairns, R. A., Harris, I. S. & Mak, T. W. Regulation of cancer cell metabolism. *Nature Rev. Cancer* **11**, 85–95 (2011).
- Pillay, K. & Govender, P. Amylin uncovered: a review on the polypeptide responsible for type II diabetes. *BioMed Res. Int.* **2013**, 826706 (2013).
- Christopoulos, G. *et al.* Multiple amylin receptors arise from receptor activity-modifying protein interaction with the calcitonin receptor gene product. *Mol. Pharmacol.* **56**, 235–242 (1999).
- Masters, S. L. *et al.* Activation of the NLRP3 inflammasome by islet amyloid polypeptide provides a mechanism for enhanced IL-1 β in type 2 diabetes. *Nature Immunol.* **11**, 897–904 (2010).
- Allen, I. C. *et al.* The NLRP3 inflammasome functions as a negative regulator of tumorigenesis during colitis-associated cancer. *J. Exp. Med.* **207**, 1045–1056 (2010).
- Vin, H. *et al.* BRAF inhibitors suppress apoptosis through off-target inhibition of JNK signaling. *eLife* **2**, e00969 (2013).

Acknowledgements We thank A. Jain, V. Pant, J. Jackson, A. Marisette, K. Michel and the Small Animal Imaging Facility (SAIF) for technical advice. This work was supported by grants to E.R.F. from NCI (R01CA160394) and (R01CA134796), CPRIT (RP120124), NCI-Cancer Center Core Grant (CA-16672) (University of Texas M.D. Anderson Cancer Center), a development award from the Lymphoma SPORE (P50CA136411), the Hildegardo E. and Olga M. Flores Foundation, and the Mel Klein Foundation and grant to J.A.B. from CPRIT (RP101243-P5). E.R.F. is a scholar of the Leukemia and Lymphoma Society, the Rita Allen Foundation and the V Foundation for Cancer Research. A.V. is a Schissler Scholar and D.C. is a CPRIT Scholar (RP101502).

Author Contributions A.V. and E.R.F. conceived the study, designed experiments and analysed data. A.V., P.R., W.N., D.C., X.S., S.K.S., M.S.R., J.L., C.V.K., E.F.S., K.N., J.P.-T., J.A.B. and K.Y.T. designed and performed experiments. P.H.G., C.C. and K.R. performed bioinformatic analyses. E.R.F. and A.V. wrote the paper. All authors discussed the paper and commented on the manuscript.

Author Information The RNA-Seq data has been deposited in the Gene Expression Omnibus (GEO) data repository and can be accessed using database accession number GSE60827. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.R.F. (elsaflares@mdanderson.org).

METHODS

Generation of $\Delta Np73$ conditional knockout mice. The Cre-loxP strategy was used to generate the $\Delta Np73$ conditional knockout allele ($\Delta Np73^{fl}$). Genomic p73 DNA from intron 3 to intron 3' was amplified from BAC clone DNA (BAC RP23-186N8, Children's Hospital Oakland Research Institute). loxP sites flanking exon 3' of p73 and neomycin (*neo*) gene flanked by frt sites inserted in intron 3' were cloned into pL253³¹. Mouse embryonic stem cells (G4) electroporated with the targeting vector were analysed by Southern blot analysis for proper targeting of the $\Delta Np73$ allele. Resulting chimaeras were mated with C57BL/6 albino females and genotyped as described below. Mice with germ line transmission of the targeted allele (conditional, flox neo allele, fn) were crossed to the FLP_{er} mice to delete the neo cassette. Resulting progeny were intercrossed with *Zp3-cre* (C57BL/6)³² transgenic mice. $\Delta Np73^{fl/+}$; *Zp3-cre* females were mated with C57BL/6 males to generate $\Delta Np73^{+/+}$ mice. The $\Delta Np73^{+/+}$ mice were intercrossed to generate $\Delta Np73^{-/-}$ mice. Compound mutant mice were generated by intercrossing the $\Delta Np63^{+/+}$ and $\Delta Np63^{fl/fl}$ (ref. 15) and the $\Delta Np73^{-/-}$ and $\Delta Np73^{fl/fl}$ mice with the $p53^{-/-}$ mice¹⁶. All procedures were approved by the IACUC at University of Texas M.D. Anderson Cancer Center.

Genotyping. Genomic DNA from tail biopsies was genotyped by Southern blot analysis by digesting genomic DNA with AflIII and HindIII or by PCR using the following primers and annealing temperatures: (1) for wild type: wt-F, 5'-ACAGT CCTCTGCTTTTACAGC-3' and wt-R (fl-R), 5'-CACACAGCACTGGCCTTGC-3', annealing temp: 58 °C, (2) for $\Delta Np73^{fl}$: fl-F, 5'-CATAGCCATGGGCTCTCTCT-3' and fl-R (wt-R), 5'-TGTCCTGCTGCTGGTTGTAT-3', annealing temp: 63 °C, (3) $\Delta Np73^{fl}$ neo: flneo-F, 5'-GGGAGGATTGGGAAGACAAT-3' and flneo-R, 5'-TGTCCTGCTGCTGGTTGTAT-3' annealing temp: 60 °C and (4) for $\Delta Np73^{KO}$: ko-F, 5'-CCTAGCCCAAGCATACTGGT-3' and wt-R, 5'-TGTCCTGCTGCTG GTTGTAT-3' annealing temp: 58 °C. Primers used to genotype for the Cre gene are as follows: Cre-F, 5'-TGGGCGGCATGGTGCAAGTT-3' and Cre-R, 5'-CGG TGCTAACACGCGTTTTC-3', annealing temp: 60 °C. The primers for $\Delta Np63^{WT}$, $\Delta Np63^{KO}$, $\Delta Np63^{fl}$ and $p53$ were previously described^{15,16}.

Cell lines. Mouse embryonic fibroblasts (MEFs) for the indicated genotypes were generated as described previously⁹. Human lung adenocarcinoma cells (H1299), colorectal adenocarcinoma cells (SW-480) and breast adenocarcinoma cells (MDA-MB-468) were purchased from ATCC and cutaneous SCC cell lines (SRB12, COLO16)³⁰ were a gift from K. Y. Tsai. The MEFs, SW-480 and MDA-MB-468 cells were cultured in DMEM (Cellgro) and H1299 cell lines were cultured in RPMI 1640 (Cellgro). The SRB12 and COLO16 cell lines were grown in DMEM/Ham's F12 50/50 (Cellgro). All cell lines used in the study tested negative for mycoplasma.

Immunohistochemistry. Mice thymic lymphomas or thymi were dissected, fixed in 10% formalin, and embedded in paraffin. Sections were de-waxed in xylene and re-hydrated using decreasing concentrations of ethanol. Antigens were unmasked in citrate buffer unmasking solution (Vector Laboratory) followed by incubation with blocking solution, and 18 h incubation at 4 °C with the following antibodies: cleaved caspase 3 (1:200) (Cell Signaling), PCNA (1:500) (Cell Signaling), malondialdehyde (1:50) (Abcam). Visualization was performed using the ImmPact DAB peroxidase substrate kit (SK4105, Vector Laboratories) and counter-stained with haematoxylin (H-3401, Vector Laboratories). The slides were mounted using VectaMount (H-5000, Vector Laboratories). Images were acquired using a Zeiss Axio microscope and analysed with ProgRes Capture Pro 4.5 software.

Senescence-associated β -galactosidase staining. Senescence-associated β -galactosidase staining on mouse thymic lymphoma was performed as described previously³³.

Quantitative real time PCR. Total RNA was prepared from MEFs or mouse tissues using TRIzol reagent (Invitrogen)^{48,34}. Complementary DNA was synthesized from 5 μ g of total RNA using the SuperScript III First-Strand Synthesis Kit (Invitrogen) according to the manufacturer's protocol followed by qRT-PCR using the SYBR Fast qPCR master mix (Kapa Biosystems). qRT-PCR was performed using a ABI 7500 Fast Real-time PCR machine. Primers for mouse *Tap63*, $\Delta Np63$, *PUMA*, *Noxa*, *bax*, *PML*, *p16* and *p21* (refs 4, 34) and human *Tap63*, $\Delta Np63$ and *GAPDH* were used as described previously^{4,34}. Human primers for *PUMA*, *Noxa*, *bax*, *PML*, *p16*, *p21* were used as described previously³³ and *GLS2* and *TIGAR* as described previously¹⁹. Mouse primers for *Tap73* are F: 5'-GCACCTACTTTGACCTCCCC-3', R: 5'-GCATGCTGAGCAAAATGAAC-3', $\Delta Np73$ are F: 5'-ATGCTTTACGT CGGTGACCC-3', R: 5'-GCATGCTGAGCAAAATGAAC-3', *IAPP* are F: 5'-C TCCAAACTGCCATCTGAGGG-3', R: 5'-CGTTTGTCCATCTGAGGGTT-3'. Human primers used for *Tap73* are F: 5'-CAGACAGCACCTACTTCGACCTT-3', R: 5'-CCGCCACACCTCATTA-3' and for $\Delta Np73$ are F: 5'-TTCAGCCAGT TGACAGAACTAAG-3', R: 5'-GGCCGTTTGTGGCATT-3'.

Western blot analysis. Fifty micrograms of protein were electrophoresed on a 10% or 15% SDS PAGE and transferred to PVDF membrane as described previously^{48,34}. Blots were probed with anti-p63 (1:500) (4A4, Santa Cruz), anti-Tap63 (1:1,000) (BioLegend), anti-Tap73 (1:500) (IMG-246, Imgenex), anti-p73 (mouse) (1:250) (IMG-259A, Imgenex), anti-p73 (1:1,000) (human) (EP436Y, Abcam), anti-p53

(WT) (1:1,000) (CM5, Vector Labs), anti-IAPP (1:1,000) (ab103580, Abcam), anti-His (1:1,000) (G18, Santa Cruz), anti-hexokinase II (1:10,000) (C64G5, Cell Signaling), anti-calcitonin receptor (1:1,000) (ab11042, Abcam), RAMP3 (1:1,000) (H125, Santa Cruz), and cleaved caspase 3 (1:1,000) (Asp 175, Cell Signaling), at 4 °C for 18 h followed by incubation for 1 h at room temperature with the appropriate secondary antibodies conjugated to horseradish peroxidase (1:5,000) (Jackson Lab). β -Actin (Sigma 1:5,000) was used as a loading control. Detection was performed using the ECL Plus Kit (Amersham) following the manufacturer's protocol and X-ray autoradiography.

Characterization of thymus using flow cytometry. Thymi from 4-week-old mice and thymic lymphomas from 10-week-old mice were collected 48 h after adenovirus infection. Single cells were obtained by homogenizing the thymi through a 0.75- μ m filter. Cells were stained with CD3-PE (145-2C11), CD4-PerCP-Cy5.5 (RM4-5), CD8-APC (53-6.7), CD45-FITC (30-F11) (BDPharmingen), AnnexinV-Pacific Blue (A35122, Life Technologies), and 7-AAD (V35124, Invitrogen) and sorted using a BD Aria Cell Sorter or analysed using the LSR Fortessa Cell Analyzer and FlowJo software.

Chromatin immunoprecipitation (ChIP). MEFs were grown to near confluence at passage 2 on DMEM media with 10% serum as previously described⁹. Thymocytes from 6-week-old mice were collected 48 h after adenovirus infection. Cellular proteins were cross-linked to DNA using 1% formaldehyde and chromatin was prepared as described previously^{48,34}. Tap63 and $\Delta Np63$ ChIP analysis was performed using a pan-p63 antibody (4A4, Santa Cruz) as described previously and the Tap73 ChIP was performed using a Tap73 antibody (ab14430, Abcam) and $\Delta Np73$ ChIP was performed using a p73 antibody (IMG 259A, Imgenex). Putative Tap63 and Tap73 binding sites were scanned 3,000 bp upstream of the 5' UTR and in intron 2 of the IAPP gene. qRT-PCR was performed by using primers specific for the indicated regions of *IAPP*: Promoter-Site 1 (−1802) forward: 5'-AGAGTTCAAGGT CATCCTCGAC-3' and (−1731) reverse: 5'-TGTTCTGACATGCAGCCTCA-3', Intron-2-Site 2 (+678) forward: 5'-AGACAGGCATGCTTAGAGACG-3' and (+765) reverse: 5'-CACTCAGTGTGGATGTCCGT-3', and non-specific site (+7532) forward: 5'-GTGTGTGATGGTTTGGTGGAT-3' and (+7623) reverse: 5'-AC AAGGCAGTTGATGGAGACT-3'. Similarly, putative $\Delta Np63$ and $\Delta Np73$ binding sites were scanned 10,000 bp upstream of the 5' UTR and in intron 1 of *Tap63* and *Tap73*. qRT-PCR was performed by using the primers specific for the indicated regions on the *Tap63* promoter: Site 1 (−41) forward: 5'-CAGGAGCTCT CAAATCAAGTCAGA-3' and (+37) reverse: 5'-ATCACAGAAGCCAGGACT TGTCAC-3', and non-specific site (−3030) forward: 5'-GCTATAAATGTTTC CATGTGATGGATTGC-3' and (−2973) reverse: 5'-TGCAGACTTAGCTATG GTCTCTTG-3'. Similarly, qRT-PCR was performed using the primers specific for the indicated regions on the *Tap73* promoter: Site 1 (−1103) forward: 5'-CTA GCACCAATCCAAGGAAAGA and (−1059) reverse: 5'-GCCTGCAGTCC GGGTTT-3' and non-specific site (−2488) forward: 5'-ACTAGACCTCTGTAC TTGTGAACATACATTT-3' and (−2382) reverse: 5'-GCACTCTCAFFATCCT GTAACAAAAA-3'.

Dual luciferase reporter assay. Luciferase assays were performed using $p53^{-/-}$; $p63^{-/-}$ and $p53^{-/-}$; $p73^{-/-}$ MEFs as described previously³⁵. To generate the luciferase reporter gene (pGL3-IAPP), the DNA fragment containing the Tap63/Tap73-binding site identified by ChIP was amplified from C57BL/6 genomic DNA by PCR with the following primers containing 5' XhoI and 3' HindIII cloning restriction enzyme sites: *IAPP* 5'-ATACTCGAGGTGTTCAGGGAACCTTCGGT-3' (forward) and 5'-ATAAAGCTTCACTGACCTCCAACTCCC-3' (reverse). Similarly, a mutant version of the luciferase reporter gene (pGL3-IAPP^{Mut}) was generated using QuikChange Lightning (Agilent Technologies) following the manufacturer's instructions. The following primers 5'-TATTGTTCTGACATCCAGCCTGATG TTGCCAGTCTGGT-3' (forward) and 5'-ACCACTGGGCAACATCAGGCTGGATGTGCAACAATA-3' (reverse) were used to generate the mutant version.

Reverse transfection. Cells were transfected with 50 nM si $\Delta Np63$ (SASI_Hs02 00328367) (Mission siRNA, Sigma), si $\Delta Np73$ (SASI_Hs02_00326884) (Mission siRNA, Sigma), siTap63 (SASI_Hs01_00246771) (Mission siRNA, Sigma), siTap73 (SASI_Hs02_00339573) (Mission siRNA, Sigma), siRAMP3 (SASI_Hs01_00199036) (Mission siRNA, Sigma), siCalcitonin receptor (SASI_Hs01_0007738) (Mission siRNA, Sigma), siIAPP (SASI_Hs01_00183962) (Mission siRNA, Sigma) or siNT (SIC_001) (Mission siRNA, Sigma) using Lipofectamine RNAiMAX (Invitrogen). The mixture of siRNA and Lipofectamine were combined together and added to the well followed by the addition of 200,000 cells per well in a six-well dish.

Transfections and generation of IAPP- and hexokinase II-expressing cells. 3×10^5 cells were plated in 10-cm dishes. MEFs and human cancer cells were transfected with 8 μ g Myc-DDK-IAPP (RC215074) (Origene) or 3.3 μ g HKII (Plasmid 25529) (Addgene) using X-tremeGENE HP (Roche) and incubated for 48–60 h. Cells were selected with G418, MEFs (350 μ g μ l^{−1}) and human cancer cells (500 μ g μ l^{−1}) for a period of 9 days.

Secreted IAPP protein concentration. Twelve hours after knockdown of $\Delta Np63/\Delta Np73$ in human cancer cells, fresh serum-free media was added to the cells. Following a sixty-hour incubation, the media was collected and concentrated using Amicon Ultra-15 Centrifugal Filter Units (UFC901008, EMD Millipore).

RNA sequencing and analysis. Five micrograms of polyA⁺ RNA were used to construct RNA-Seq libraries using the standard Illumina protocol. Mouse mRNA sequencing yielded 30–40 million read pairs for each sample. The mouse mRNA-Seq reads were mapped using TopHat³⁶ onto the mouse genome and build UCSC mm9 (NCBI 37) and the RefSeq mouse genes. Gene expression and gene expression differences were computed using Cufflinks³⁶. For each species, a combined profile of all samples was computed; mRNA abundance was mean-centred and Z-score transformed for each mRNA individually. Principal component analysis was executed using the implementation within the R statistical analysis system. Hierarchical clustering of samples was executed by first computing the symmetrical sample distance matrix using the Pearson correlation between mRNA profiles as a metric, supervised sample analysis was performed using the *t*-test statistics, and heat maps were generated using the heatmap.2 package in R. For gene signatures and pathway analysis gene list from the RNA-Seq comparing $\Delta Np63^{\Delta/\Delta};p53^{-/-}$ versus $\Delta Np63^{\Delta/\Delta};p53^{-/-}$ and $\Delta Np73^{\Delta/\Delta};p53^{-/-}$ were obtained at a *P* value < 0.01. The genes upregulated in the $\Delta Np63^{\Delta/\Delta};p53^{-/-}$ and $\Delta Np73^{\Delta/\Delta};p53^{-/-}$ and down regulated in the $\Delta Np63^{\Delta/\Delta};p53^{-/-}$ were selected. The relative fold change of the genes were calculated and sorted from highest to lowest. Genes with a greater than 1.5-fold-increase were selected and run through the ingenuity pathway analysis (IPA) (Ingenuity Systems) to screen for pathways and processes. Genes from the selected pathways were cross-referenced with the Gene Set Enrichment (GSEA) (Broad Institute) data analysis, DAVID Bioinformatics Resource 6.7 and GSEA implementation at the Molecular Signature Database (MSigDB)³⁷.

Magnetic resonance imaging. MRI imaging was performed at 10 weeks of age when the tumours were established and the volumes range from 2.3 mm³ to 5 mm³. To reduce the variation between different groups of mice, a cohort of *n* = 5 with similar tumour volumes was established and tumours regression was monitored by MRI. All mice were scanned once a week for a period of 35 weeks on a 7-T, 30-cm bore BioSpec MRI system (Bruker Biospin Corp., Billerica, MA).

Hyperpolarized magnetic resonance spectroscopy. Dynamic MR spectroscopy (MRS) of hyperpolarized (HP) [1-¹³C]pyruvate was performed *in vivo* in tumour-bearing mice. To achieve polarization, a 26-mg sample of pyruvic acid (Sigma-Aldrich, St. Louis, MO) with 15 mM of OX063 radical (GE Healthcare, Waukesha, WI) and 1.5 mM Prohance (Bracco Diagnostics Inc., Monroe Township, NJ) was polarized in a HyperSense DNP system (Oxford Instruments, Abington, Oxfordshire, UK) as previously described^{38,39}. The frozen sample was dissolved in a 4 ml buffer containing 40 mM Tris, 80 mM NaOH, and 50 mM NaCl, resulting in a final isotonic and neutral solution containing 80 mM [1-¹³C]pyruvate. A dual-tuned ¹H/¹³C linear radio frequency volume coil with 72 mm internal diameter (ID) was used in conjunction with imaging gradients with 12 cm ID. For anatomic imaging, the ¹H channel was used in transmit/receive mode. In addition to localizing scans, flow-weighted oblique gradient echo images (TE = 1.4 ms; TR = 55 ms; 90° excitation; 3 cm × 3 cm field-of-view (FOV) encoded over a 64 × 64 image matrix) were acquired to confirm that the slice prescription for ¹³C measurements would not be obscured by signals originating from within the heart. For carbon spectroscopy, the radio frequency volume coil was used in transmit-only mode in conjunction with a custom-built 15-mm ID ¹³C surface coil for signal reception. After dissolution, 200 µl of the HP [1-¹³C]pyruvate solution was administered to the animals via tail-vein catheter. A slice-selective pulse-acquire sequence (TR = 1,500 ms; 15° flip angle; 5 kHz spectral bandwidth; 2,048 spectral points; 8 mm oblique slab; 120 repetitions) was used for dynamic spectroscopy beginning approximately 15 s before injection. Data were processed to generate spectral time-courses of the HP-pyruvate and its lactate product. Spectra were phase adjusted and the area under the spectral peaks associated with [1-¹³C]pyruvate and [1-¹³C] lactate were integrated over time to reflect the overall signal observed from each metabolite over the course of the measurement. Total lactate signal, which could only arise from interaction of HP pyruvate with relevant metabolic enzymes, was normalized to the total signal from pyruvate.

Glycolysis stress assay. Extra-cellular acidification rate (ECAR) was measured using the extracellular flux analyser (SeaHorse Bioscience XF96) following the manufacturer's instructions. Forty-eight hours after transfection, the cells were plated at a density of 1.5×10^4 cells per well in the XF 96-well cell culture plates. Twenty-four hours after seeding, the culture medium was replaced with 180 µl of running medium and incubated for 1 h at 37 °C in a non-CO₂ incubator. Before calibration, 20 µl of 50 mM glucose, 11 µM oligomycin and 650 mM 2DG were aliquoted into each port in the sensor cartridge. ECAR was measured after the addition of glucose and oligomycin and before the addition of 2DG. Extra-cellular acidification rate was normalized to mpH min⁻¹.

Glucose uptake measurement. Glucose uptake was calculated as a measure of glucose-dependent proton secretion from the maximum and basal glucose consumption

after addition of 20 µl of 50 mM glucose and measured using the extracellular flux analyser (SeaHorse Biosciences XF96).

Glucose-6-phosphate assay. Glucose-6-phosphate was measured using a glucose-6-phosphate assay kit (ab83426, Abcam) following the manufacturer's instructions. Forty-eight hours after transfection, 2×10^6 cells were collected, homogenized and passed through a 10-kDa spin-column filter. The eluate was collected and glucose-6-phosphate enzyme and substrate reaction was performed for 30 min and absorbance was measured at 450 nm.

Proliferation assay. The transfected human cancer cells were plated at a density of 5×10^3 cells in 6 replicates in a 96-well dish. Twelve hours later, the cells were labelled with 10 mM EdU (5'-ethynyl-2'-deoxyuridine) for a period of 8 h. The assay was performed using the Click-iT EdU microplate assay (Invitrogen). Images were obtained using a Zeiss Axio fluorescent microscope and analysed using the AxioVision Image 4.5 software.

Apoptosis assay. Cells were plated at a density of 1×10^4 cells in 6 replicates in a 96-well dish. Twelve hours later, the cells were washed with $1 \times$ annexin-binding buffer and a cocktail of 5 µl annexin V-Alexa Fluor 488 for $100 \mu\text{g ml}^{-1}$ propidium iodide (PI) and $2 \mu\text{g ml}^{-1}$ Hoechst 33342 (Invitrogen) was added. Images were captured using the Zeiss fluorescent microscope and Axiovision Image 4.5 software. Quantification of the percent apoptosis was obtained using a high-throughput immunofluorescence plate reader (Celigo).

ROS assay. Cells were plated at a density of 1×10^4 cells in 6 replicates in 96-well dish. Twelve hours later, the cells were incubated with a cocktail of 5 µM concentration of CellROX Deep Red Reagent (C10422, Invitrogen) and $2 \mu\text{g ml}^{-1}$ Hoechst 33342 (Invitrogen) for 45 min at 37 °C. Images were captured using a Zeiss fluorescent microscope and Axiovision Image 4.5 software. Quantification of the percent ROS was obtained using a high-throughput immunofluorescence plate reader (Celigo)⁴⁰.

In vitro adeno-Cre infection. $\Delta Np63^{\Delta/\Delta};p53^{-/-}$ and $\Delta Np73^{\Delta/\Delta};p53^{-/-}$ MEFs were plated at a density of 2.5×10^5 cells in 10-cm dishes before infection. Twelve hours later, MEFs were infected with Adeno-CMV-mCherry or Adeno-CMV-Cre-mCherry (Gene Transfer Vector Core Facility, University of Iowa). The cells were infected at a multiplicity of infection of 6,000 particles per cell. The efficiency of infection was quantified by assessing mCherry-positive cells.

In vivo adeno-virus infection and IVIS Lumina imaging. All mice were anesthetized using isoflurane and 2% oxygen and placed on a custom bed. An incision was performed to expose the sternum. Using a 28.5G U100 Insulin syringe, Adeno-mCherry/Adeno-Cre-mCherry (Gene Transfer Vector Core Facility, University of Iowa), Adeno-IAPP-mCherry (Vector Labs) or Adeno-shIAPP-U6-mCherry (TRCN0000416196, Mission shRNA) (Vector Labs) (sequence CCGGTGTAATTCATGCTAAGAACTCGAGTCTTAGCATGAGAAATTTACATTTTGTG) was surgically administered by intra-thymic injection (5×10^{12} viral particles per gram of body weight) through the 2nd and 3rd sternum. The incision was sealed using wound clips and mice were allowed to recover. To determine the efficiency of the *in vivo* viral delivery to the thymic lymphoma, IVIS Lumina Imaging (Perkin Elmer) was performed 48 h later. Images were captured using a Mid-600 series bandwidth filter and analysed using the Living Image data analysis software.

shRNA knockdown. shRNA plasmids for Trp63 (Clone ID: V3LMM_508694) (sequence TGATCTTCAGCAACATCTC) and Trp73 (Clone ID: V3LMM_438557) (sequence TGCAGGTGGAAGACATCCA) were obtained from the MD Anderson shRNA core facility (Open Biosystems). 293T cells were plated at a density of 2.5×10^5 cells in 10 cm dishes. Three micrograms of shRNA and packaging vectors were transfected as described previously⁴. Cells were selected using puromycin ($3 \mu\text{g ml}^{-1}$) for 7 days.

In vitro and in vivo administration of 2-deoxy-D-glucose. 1×10^4 cells were plated in 6 replicate wells in a 96-well dish. Twelve hours later, the human cancer cells were treated with 50 mM final concentration of 2-deoxy-D-glucose (2DG) (D8375-5G, Sigma) for 1 h. Similarly, 2DG (500 mg per kg of tumour weight) (D8375-5G-Sigma) was administered directly into the lymphoma of mice as described earlier³⁹.

N-acetyl-L-cysteine treatment. 1×10^4 cells were plated in 6 replicate wells in a 96-well dish. Twelve hours later, cells were treated with N-acetyl-L-cysteine (NAC) (2 mM) (A8199, Sigma) final concentration for a period of 1 h.

Amylin and caspase inhibitor treatment. 2×10^5 cells were plated in triplicate in a 6-well dish. Twelve hours later, cells were treated with Amylin peptide (5 µM) (A5972, Sigma) or with a caspase 1 inhibitor (20 µM) (Z-YVAD-FMK-218746, Calbiochem) for a period of 48 h.

In vitro and in vivo administration of pramlintide acetate. 2×10^5 cells were plated in duplicate in a 6-well dish. Twelve hours later, cells were treated with $10 \mu\text{g ml}^{-1}$ pramlintide acetate (AMYLIN Pharmaceuticals) or placebo for a period of 48 h. pramlintide acetate (AMYLIN Pharmaceuticals) or placebo (sodium acetate/acetic acid) was surgically administered through non-invasive intra-thymic injection using a multiple dose protocol of pramlintide acetate (30 µg per gram of tumour weight). One injection per week for three weeks was administered directly into the thymic

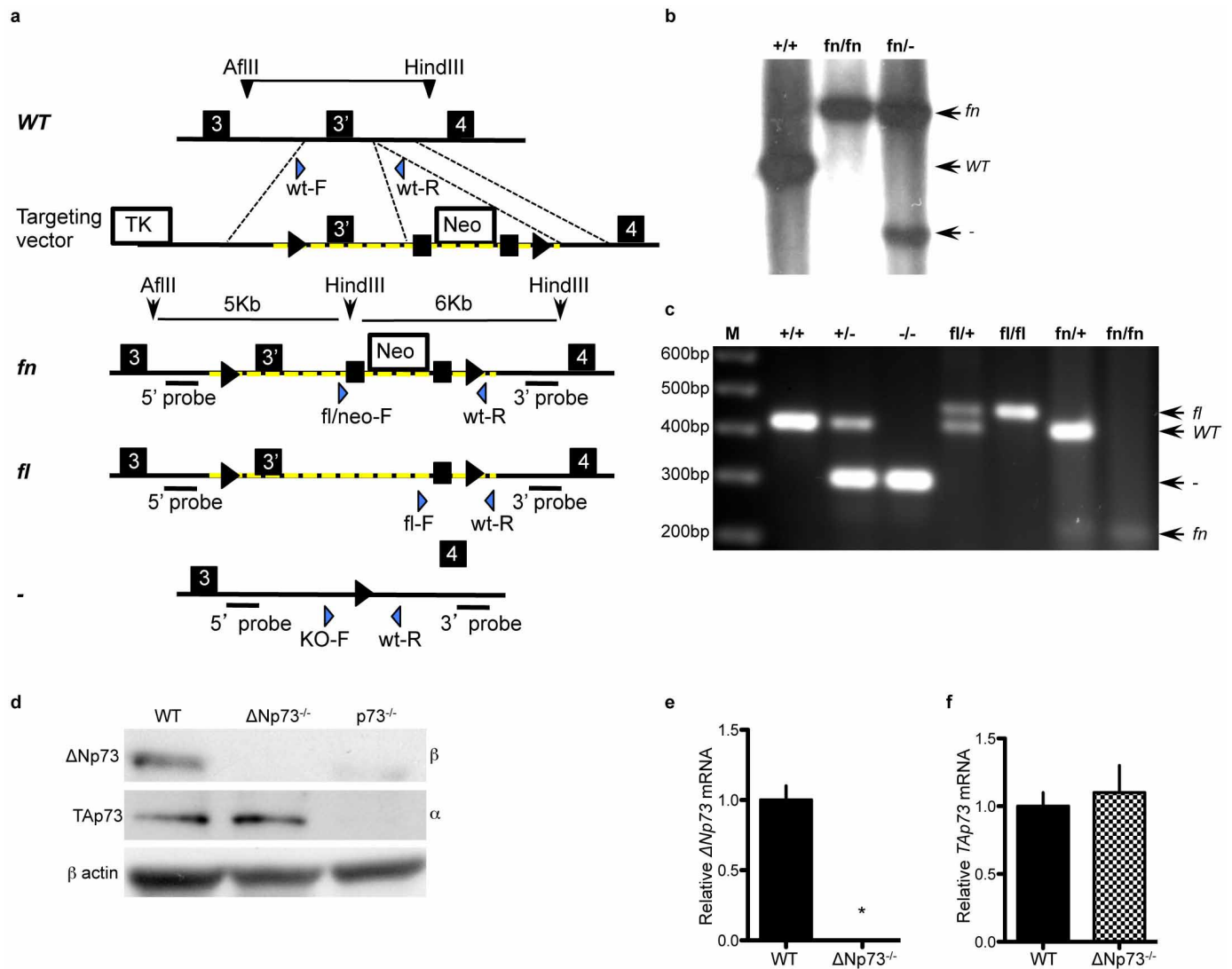
lymphoma of the animal. Another cohort of mice was treated bi-weekly for 3 weeks by intra-venous (IV) tail-vein injection of pramlintide acetate (45 µg per kg body weight) or placebo. The investigator was blinded to the treatment administered to each mouse. Tumour volumes were monitored weekly by MRI. Health and blood glucose levels of the treated animals were monitored weekly.

In vitro and in vivo administration of calcitonin receptor antagonist. 2×10^5 cells were plated in duplicate in a 6-well dish. Twelve hours later, cells were treated with Calcitonin receptor antagonist (1 nM) (AC187, Tocris Bioscience) for a period of 48 h with or without simultaneous pramlintide treatment. Similarly, a chronic dose of calcitonin receptor antagonist (1 nmol per gram of tumour weight) was administered through non-invasive intra-thymic injections with one injection every week for a period of three weeks with or without simultaneous pramlintide treatment. Tumour volume was monitored and measured weekly by MRI.

Survival analysis. Survival analysis was conducted for the IAPP, RAMP3 and CalCR gene in the following data sets: the Memorial Sloan Kettering Cancer Center and the TCGA Cancer cohort. We considered four major cancer types with high p53 mutation rates, which include lung squamous cell carcinoma⁴¹, head and neck squamous cell cancer^{42,43}, basal breast cancer^{44,45}, and colon cancer⁴⁶. The co-expression of the three genes was analysed in cases only with p53 mutation. In all cases, we considered gene expression changes above or below two standard deviations with respect to the normal controls. The log-rank test and Cox *P* test was used to assess significance between the samples with or without expression changes of the IAPP, RAMP3 and CalCR gene using the cBioPortal for cancer genomics⁴⁷.

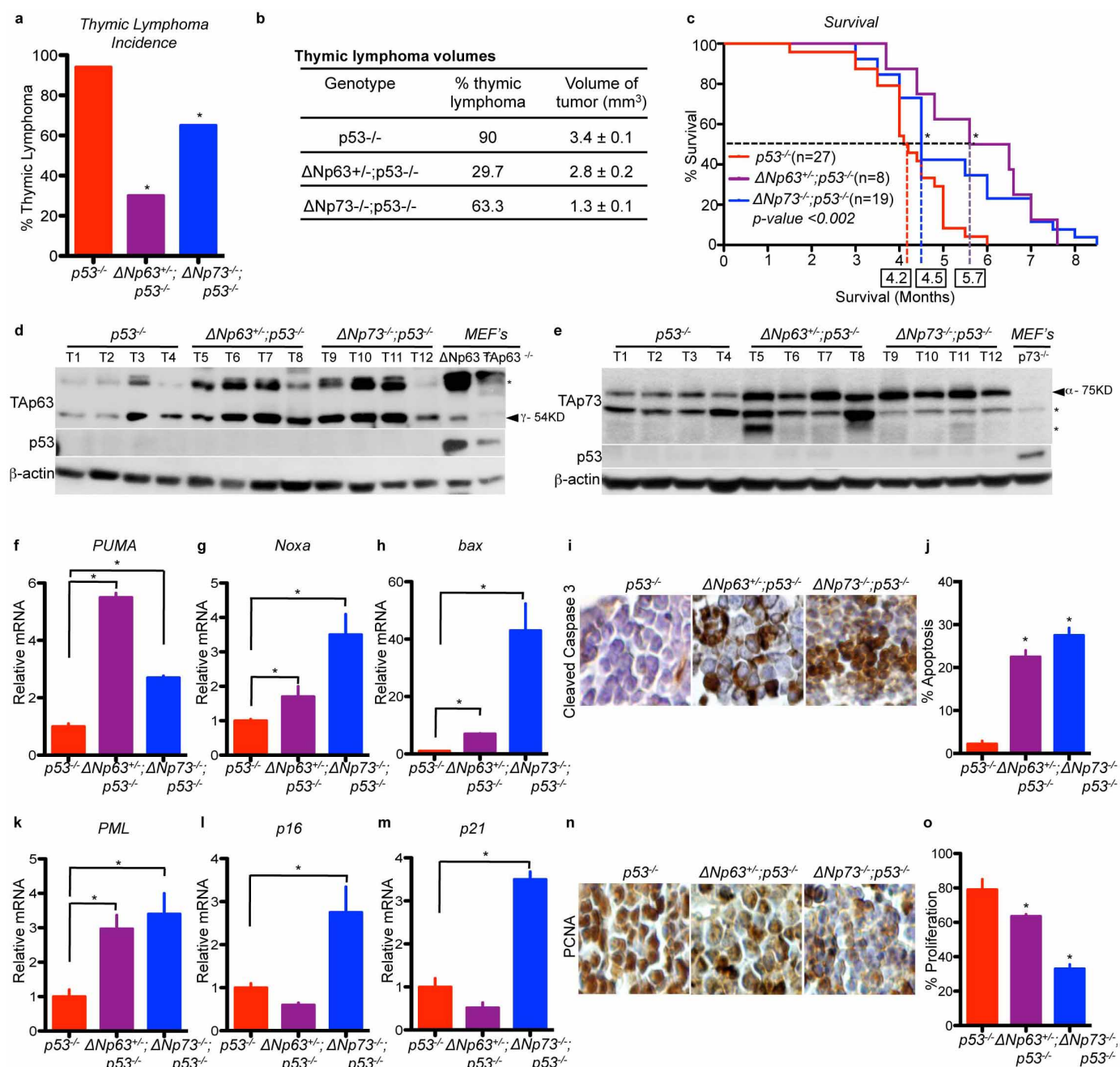
Statistics. Sample size for mouse cohorts in each experiment was chosen based on the penetrance of the thymic lymphoma phenotype of the *p53*^{-/-} mouse model (80%). Twenty to thirty mice were used for survival analyses. Data were analysed using a one-way ANOVA test or a Student's *t*-test (two-sided) was used for comparison between two groups of data. A *P* value of 0.05 was considered significant. Data are represented as mean \pm s.e.m.

31. Liu, P., Jenkins, N. A. & Copeland, N. G. A highly efficient recombineering-based method for generating conditional knockout mutations. *Genome Res.* **13**, 476–484 (2003).
32. Lewandoski, M., Wassarman, K. M. & Martin, G. R. *Zp3-cre*, a transgenic mouse line for the activation or inactivation of *loxP*-flanked target genes specifically in the female germ line. *Curr. Biol.* **7**, 148–151 (1997).
33. Jackson, J. G. *et al.* p53-mediated senescence impairs the apoptotic response to chemotherapy and clinical outcome in breast cancer. *Cancer Cell* **21**, 793–806 (2012).
34. Su, X. *et al.* TAp63 prevents premature aging by promoting adult stem cell maintenance. *Cell Stem Cell* **5**, 64–75 (2009).
35. Lin, Y. L. *et al.* p63 and p73 transcriptionally regulate genes involved in DNA repair. *PLoS Genet.* **5**, e1000680 (2009).
36. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnol.* **28**, 511–515 (2010).
37. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**, 44–57 (2009).
38. Ardenkjaer-Larsen, J. H. *et al.* Increase in signal-to-noise ratio of > 10,000 times in liquid-state NMR. *Proc. Natl Acad. Sci. USA* **100**, 10158–10163 (2003).
39. Sandulache, V. C. *et al.* Glycolytic inhibition alters anaplastic thyroid carcinoma tumor metabolism and improves response to conventional chemotherapy and radiation. *Mol. Cancer Ther.* **11**, 1373–1380 (2012).
40. Maddocks, O. D. *et al.* Serine starvation induces stress and p53-dependent metabolic remodelling in cancer cells. *Nature* **493**, 542–546 (2013).
41. The Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525 (2012).
42. Agrawal, N. *et al.* Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in *NOTCH1*. *Science* **333**, 1154–1157 (2011).
43. Stransky, N. *et al.* The mutational landscape of head and neck squamous cell carcinoma. *Science* **333**, 1157–1160 (2011).
44. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
45. Banerji, S. *et al.* Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature* **486**, 405–409 (2012).
46. The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
47. Cerami, E. *et al.* The cBio Cancer Genomics Portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discovery* **2**, 401–404 (2012).



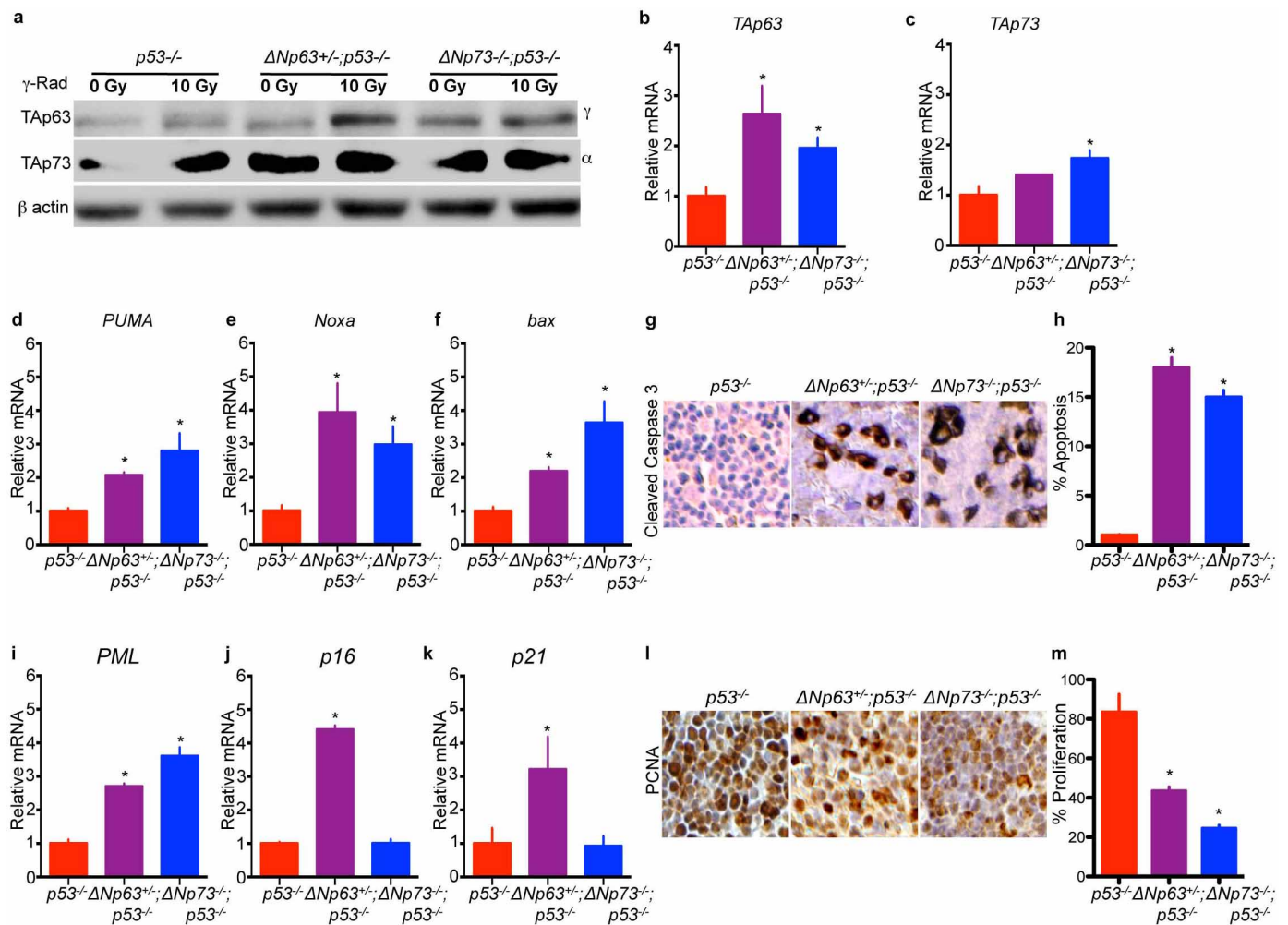
Extended Data Figure 1 | Generation and characterization of $\Delta Np73$ conditional knockout mice. **a**, The $\Delta Np73$ targeting vector was generated by inserting *loxP* sites (triangles) flanking exon 3' and a neomycin cassette (neo) flanked by *frt* sites (squares). The location of PCR primers in each allele is shown by blue arrows. The targeted region of the floxed allele is depicted by yellow-dashed lines. **b**, Southern blot analysis using the 5' probe shown in **a** and

tail genomic DNA derived from mice of the indicated genotypes. **c**, PCR analysis using tail genomic DNA of the indicated genotypes. **d**, Western blot analysis using mouse embryo fibroblasts (MEFs) of the indicated genotypes. **e**, **f**, qRT-PCR in MEFs of the indicated genotypes, $n = 4$, $P < 0.005$. Statistical significance is indicated by black asterisks.



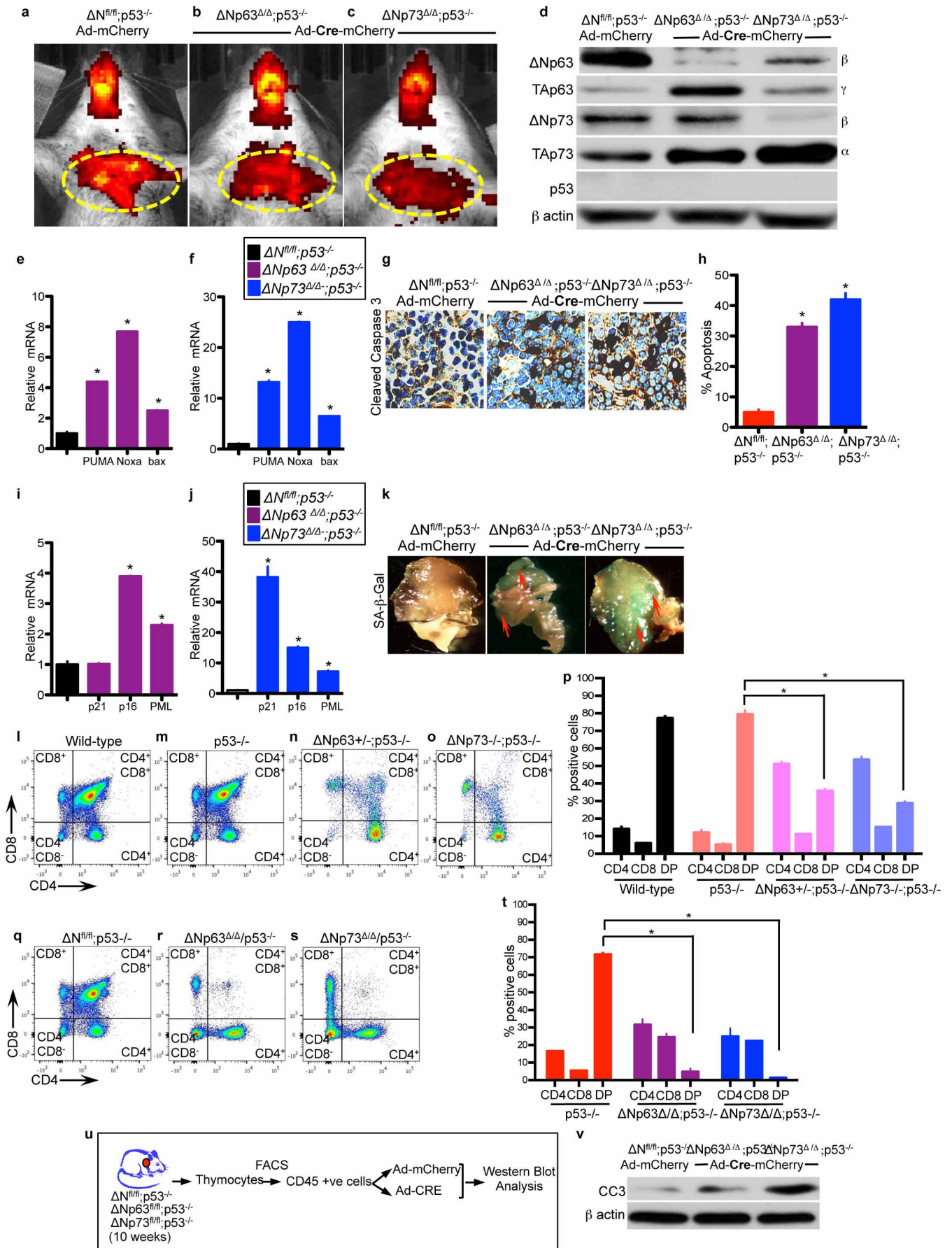
Extended Data Figure 2 | Decreased thymic lymphomagenesis and increased survival in mice double deficient for *ΔNp63* and *p53* or *ΔNp73* and *p53*. **a**, Quantification of thymic lymphoma incidence ($n = 30$ mice). **b**, Table showing thymic lymphoma volumes. The difference in tumour volumes between *p53*^{-/-} and *ΔNp63*^{+/-};*p53*^{-/-} and *p53*^{-/-} and *ΔNp73*^{+/-};*p53*^{-/-} was statistically significant with P values < 0.03 and < 0.002 , respectively. **c**, Kaplan–Meier survival in mice. Boxed numbers indicate median survival. **d**, **e**, Western blot analysis of thymic lymphomas of the indicated genotypes. Arrows indicate specific isoforms, and asterisks

indicate non-specific bands. **f–h**, qRT–PCR for *PUMA* (**f**), *Noxa* (**g**), and *bax* (**h**) in thymic lymphomas of the indicated genotypes, $n = 4$, $P < 0.005$. **i**, Immunohistochemistry (IHC) for cleaved caspase 3 in thymic lymphomas. **j**, Quantification of apoptosis as assessed by cleaved caspase 3 staining, $n = 20$ fields of 3 biological replicates, $P < 0.005$. **k–m**, qRT–PCR for *PML* (**k**), *p16* (**l**), and *p21* (**m**) in indicated thymic lymphomas, $n = 4$, $P < 0.005$. **n**, IHC for PCNA in indicated thymic lymphomas. **o**, Quantification of the percentage of proliferation as assessed by PCNA staining, $n = 20$ fields of 3 biological replicates, $P < 0.005$. Statistical significance indicated by black asterisks.



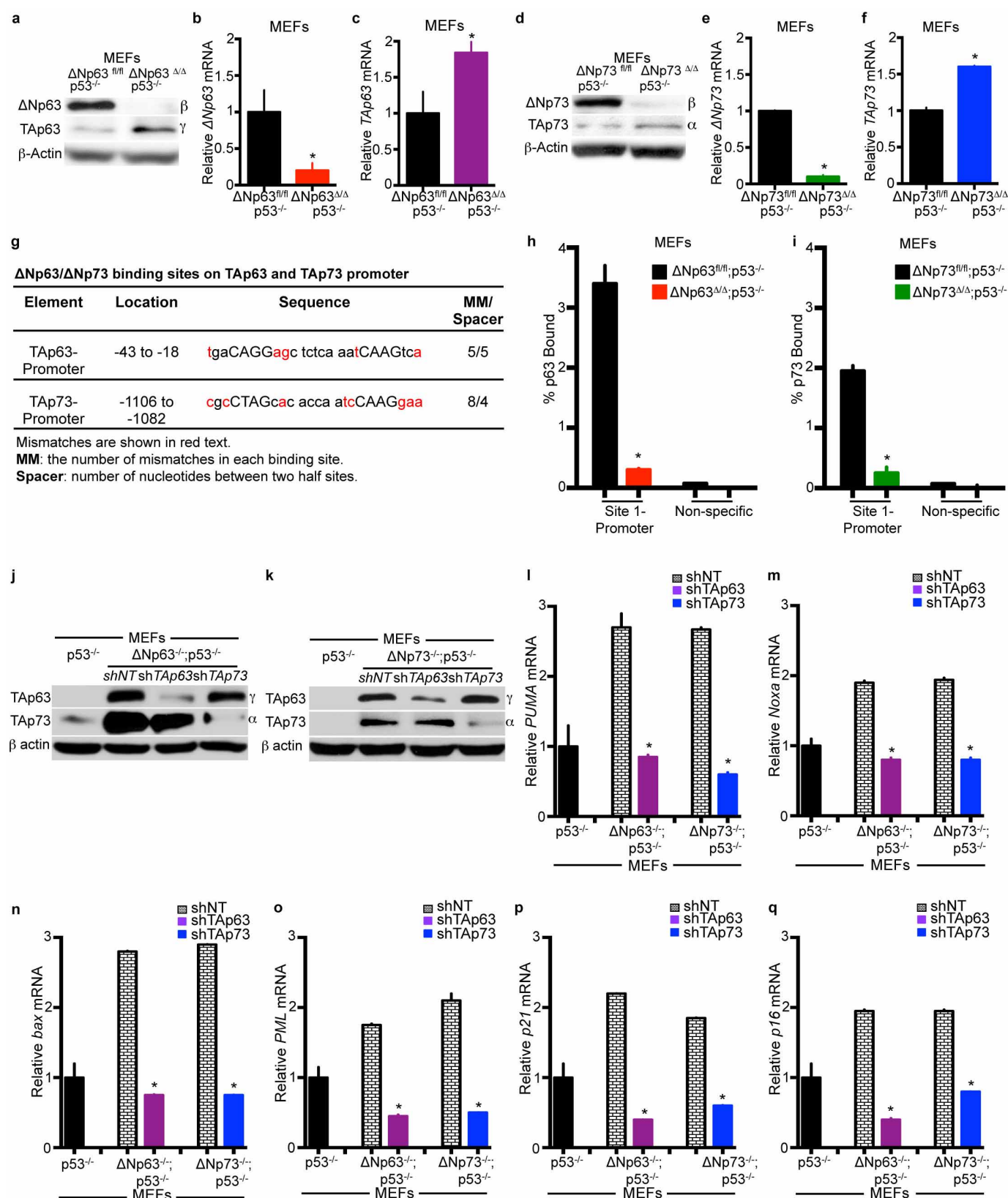
Extended Data Figure 3 | Increased apoptosis and cell cycle arrest in $\Delta Np63^{+/-};p53^{-/-}$ and $\Delta Np73^{-/-};p53^{-/-}$ thymocytes after genotoxic stress. **a**, Western blot analysis in thymocytes derived from mice 6 h after treatment with 0 Gy or 10 Gy gamma irradiation. **b–f**, qRT-PCR for TAp63 (**b**), TAp73 (**c**), PUMA (**d**), Noxa (**e**), and bax (**f**) from samples shown in **a**, $n = 4$, $P < 0.005$. qRT-PCR normalized to samples treated with 0 Gy. **g**, Immunohistochemistry (IHC) for cleaved caspase 3 in samples from **a**.

h, Quantification of the percentage of apoptosis as assessed by cleaved caspase 3 staining, $n = 20$ fields of 3 biological replicates, $P < 0.005$. **i–k**, qRT-PCR for PML (**i**), p16 (**j**), and p21 (**k**) using total RNA from samples shown in **a**, $n = 4$, $P < 0.005$. **l**, IHC for PCNA in samples shown in **a**. **m**, Quantification of the percentage of proliferation as assessed by PCNA staining, $n = 20$ fields of 3 biological replicates, $P < 0.005$. Statistical significance is indicated by black asterisks.



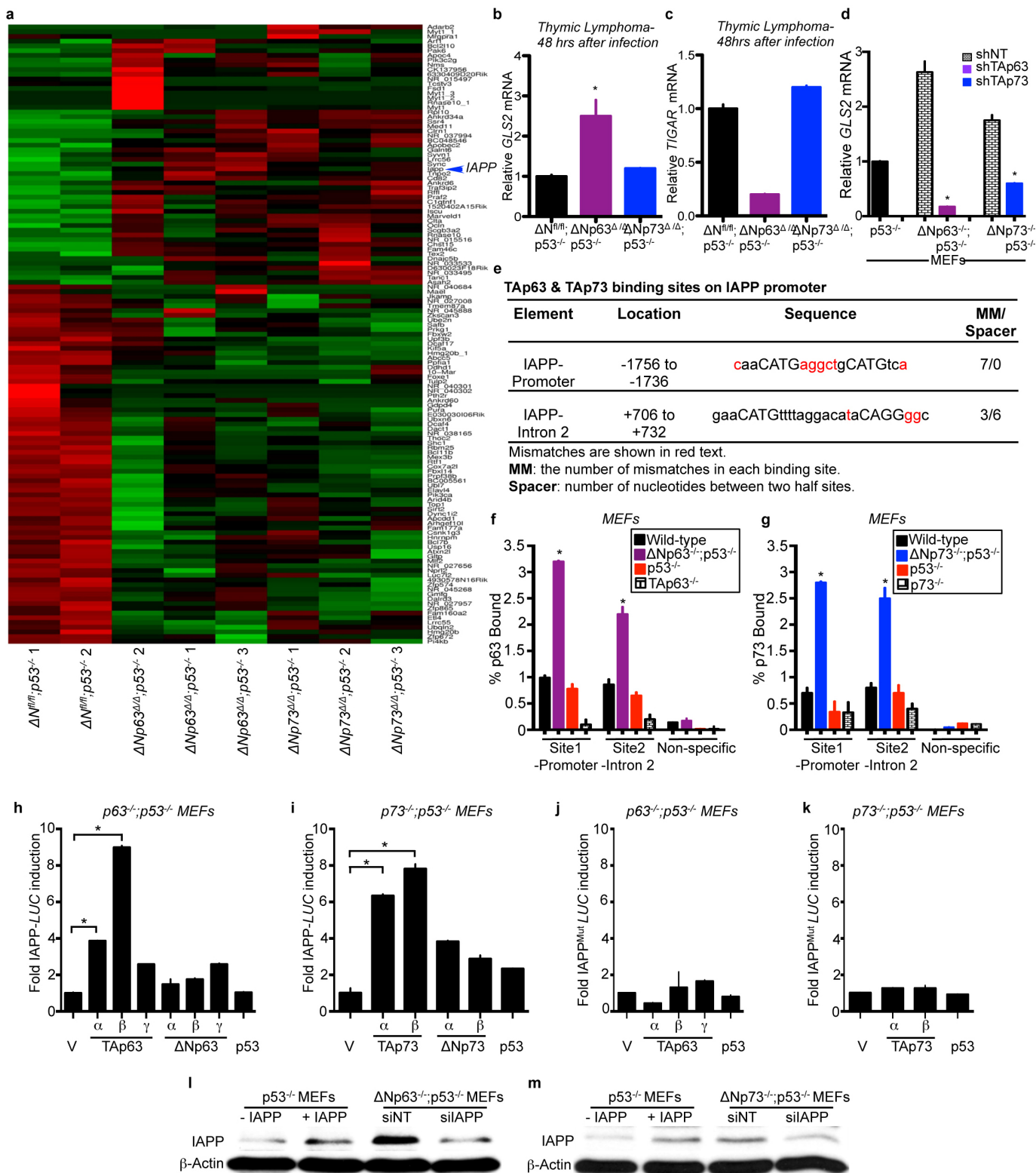
Extended Data Figure 4 | *In vivo* intra-thymic delivery of adenovirus-Cre-mCherry. **a–c**, IVIS Lumina imaging of thymic lymphomas of mice of the indicated genotypes infected with adenovirus (Ad)-mCherry (**a**) or Ad-Cre-mCherry (**b, c**) at 10 weeks of age and 48 h after adenoviral delivery. Red fluorescence indicates viral delivery to the thymus shown by the yellow dashed ovals. Red fluorescence near the mouth is due to auto-fluorescence of calcium and mineral deposits in the teeth. **d**, Western blot analysis using lysates from indicated thymic lymphomas 48 h after infection with adenovirus (Ad)-mCherry or Ad-Cre-mCherry. **e, f**, Quantitative real time (qRT-PCR) of thymic lymphomas 48 h after infection with Ad-mCherry ($\Delta N^{\beta/\beta}; p53^{-/-}$) or Ad-Cre-mCherry ($\Delta Np63^{\Delta/\Delta}; p53^{-/-}$ or $\Delta Np73^{\Delta/\Delta}; p53^{-/-}$), $n = 4$, $P < 0.005$. **g**, Immunohistochemistry (IHC) for cleaved caspase 3 in thymic lymphomas 48 h after infection with Ad-mCherry ($\Delta N^{\beta/\beta}; p53^{-/-}$) or Ad-Cre-mCherry ($\Delta Np63^{\Delta/\Delta}; p53^{-/-}$ or $\Delta Np73^{\Delta/\Delta}; p53^{-/-}$). **h**, Quantification of apoptosis as assessed by cleaved caspase 3 staining of the indicated thymic lymphomas, $n = 20$ fields of 3 biological replicates, $P < 0.005$. **i, j**, qRT-PCR of thymic lymphomas 48 h after treatment with Ad-mCherry ($\Delta N^{\beta/\beta}; p53^{-/-}$) or

Ad-Cre-mCherry ($\Delta Np63^{\Delta/\Delta}; p53^{-/-}$ or $\Delta Np73^{\Delta/\Delta}; p53^{-/-}$), $n = 4$, $P < 0.005$. **k**, Senescence-associated β -galactosidase (SA- β -gal) staining (blue) of thymic lymphomas 48 h after treatment with Ad-mCherry ($\Delta N^{\beta/\beta}; p53^{-/-}$) or Ad-Cre-mCherry ($\Delta Np63^{\Delta/\Delta}; p53^{-/-}$ or $\Delta Np73^{\Delta/\Delta}; p53^{-/-}$). **l–o**, Flow cytometry plots of the indicated thymocytes at 4-week of age. **p**, Bar graph showing quantification of CD4, CD8, and CD4/CD8 double-positive (DP) cells. $n = 3$ mice per genotype, $P < 0.005$. **q–s**, Flow cytometry plots of thymic lymphoma cells 48 h after adenovirus-mCherry or adenovirus-CRE treatment for the indicated genotypes. **t**, Bar graph showing quantification of CD4, CD8, and CD4/CD8 double-positive (DP) cells in the indicated genotypes. $n = 3$ mice per genotype, $P < 0.005$. **u**, Cartoon representation of isolation of CD45-positive thymic lymphoma cells from 10-week-old mice of indicated genotypes. **v**, Western blot analysis of CD45-positive thymic lymphoma cells after treatment with Ad-mCherry ($\Delta N^{\beta/\beta}; p53^{-/-}$) or Ad-CRE-mCherry ($\Delta Np63^{\Delta/\Delta}; p53^{-/-}$ and $\Delta Np73^{\Delta/\Delta}; p53^{-/-}$). Statistical significance is indicated by black asterisks.



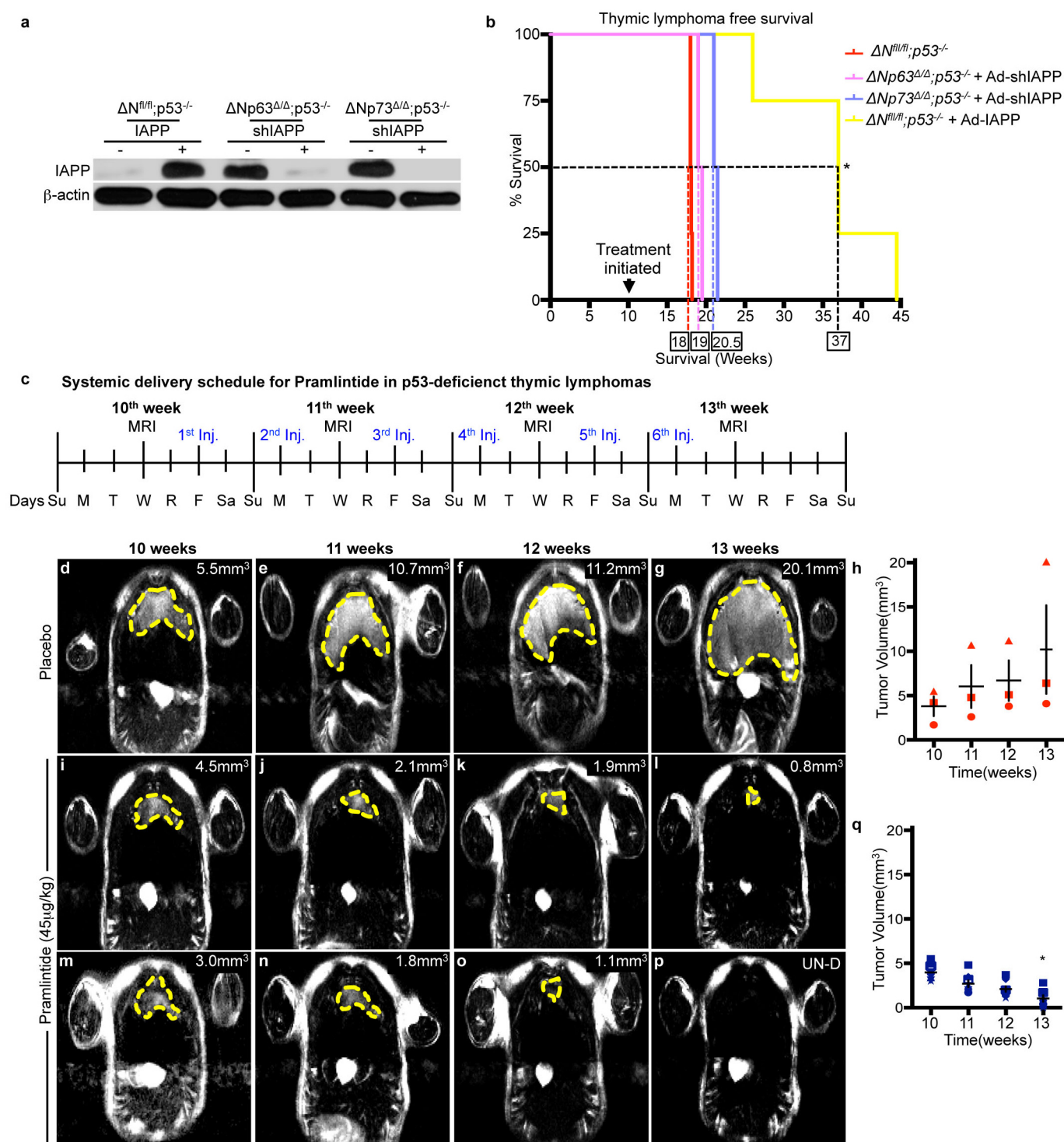
Extended Data Figure 5 | Loss of $\Delta Np63/\Delta Np73$ induces TAp63 and TAp73 upregulation in the absence of p53. **a**, Western blot analysis in $\Delta Np63^{fl/fl}; p53^{-/-}$ MEFs before ($\Delta Np63^{fl/fl}; p53^{-/-}$) and after ($\Delta Np63^{\Delta/\Delta}; p53^{-/-}$) Ad-Cre administration. **b**, **c**, qRT-PCR for $\Delta Np63$ (**b**) and TAp63 (**c**) in indicated MEFs. **d**, Western blot analysis in $\Delta Np73^{fl/fl}; p53^{-/-}$ and $\Delta Np73^{\Delta/\Delta}; p53^{-/-}$ MEFs. **e**, **f**, qRT-PCR for $\Delta Np73$ (**e**) and TAp73 (**f**) in indicated MEFs, $n = 4$, $P < 0.005$. **g**, Table showing $\Delta Np63$ and $\Delta Np73$ binding sites on the TAp63 and TAp73 promoter regions. **h**, **i**, qRT-PCR of

chromatin immunoprecipitation using indicated MEFs and an antibody for p63 (**h**) or p73 (**i**) $n = 3$, $P < 0.005$. **j**, **k**, Western blot analysis in $\Delta Np63^{-/-}; p53^{-/-}$ (**j**) or $\Delta Np73^{-/-}; p53^{-/-}$ (**k**) MEFs treated with the indicated shRNAs; (shNT) indicates a non-targeting scramble shRNA. **l**–**q**, qRT-PCR for PUMA (**l**), Noxa (**m**), bax (**n**), PML (**o**), p21 (**p**), and p16 (**q**) in the indicated MEFs expressing the indicated shRNAs, $n = 5$, $P < 0.005$. Statistical significance indicated by black asterisks.



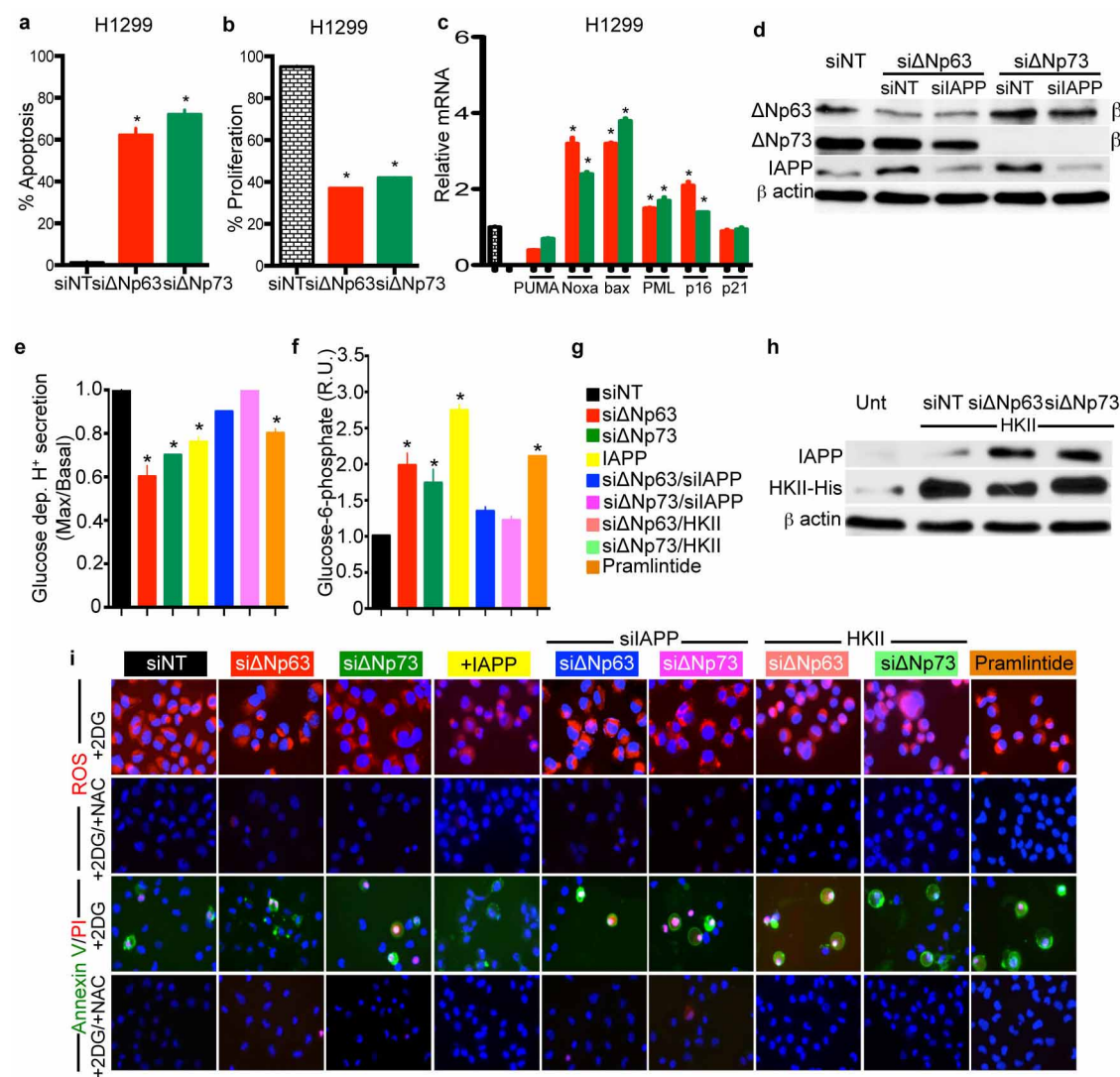
Extended Data Figure 6 | Metabolic genes including *IAPP* are upregulated in thymic lymphomas deficient for Δ Np63 or Δ Np73 and p53. **a**, Supervised hierarchical clustering of RNA-sequencing data from thymic lymphomas 48 h after treatment with Ad-mCherry (Δ N^{n/n}, p53^{-/-}) or Ad-Cre-mCherry (Δ Np63 Δ /p53^{-/-} or Δ Np73 Δ /p53^{-/-}). **b, c**, qRT-PCR for *GLS2* (**b**) and *TIGAR* (**c**) in the indicated thymic lymphomas, $n = 4$, $P < 0.005$. **d**, qRT-PCR for *GLS2* in MEFs of the indicated genotypes expressing shRNAs for a non-targeting sequence (shNT), *Tap63* (shTap63) and *Tap73* (shTap73), $n = 4$, $P < 0.005$. **e**, Table showing the Tap63 and Tap73 binding sites on the

IAPP promoter and intron 2. **f, g**, qRT-PCR of promoter site 1 using chromatin immunoprecipitation in MEFs of the indicated genotypes, $n = 3$, $P < 0.005$. **h-k**, Dual luciferase reporter assay for pGL3-*IAPP*-promoter site 1 (**h**, **i**) and a mutant version of this reporter gene (pGL3-*IAPP* MUT) (**j**, **k**). Genotypes of MEFs and vectors used are shown. V represents pcDNA3 vector. **l, m**, Western blot analysis of the indicated MEFs expressing IAPP or siRNAs for a non targeting sequence (siNT) or IAPP (siIAPP). Statistical significance indicated by black asterisks.



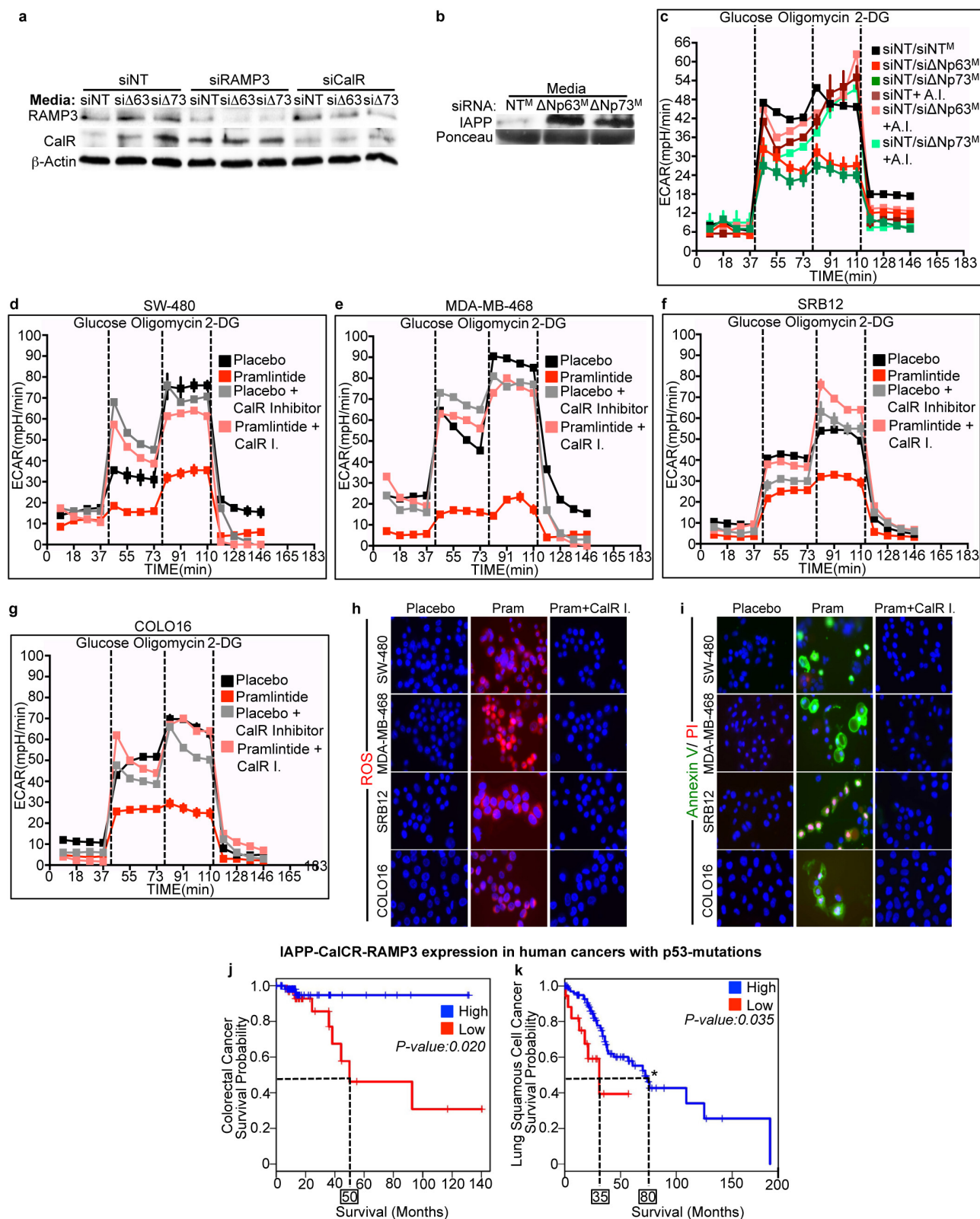
Extended Data Figure 7 | Systemic *in vivo* delivery of pramlintide results in tumour regression in p53-deficient thymic lymphomas. **a**, Western blot analysis showing IAPP expression in the indicated thymic lymphomas, $n = 5$ mice. **b**, Kaplan-Meier survival indicating thymic lymphoma-free survival. $n = 8$ mice per group, $P < 0.005$. **c**, Cartoon indicating schedule of MRI

imaging and injection (Inj.) of pramlintide in mice with p53-deficient thymic lymphomas. **d-q**, MRI imaging at 10, 11, 12 and 13 weeks after treatment with placebo (**d-g**) or pramlintide (**i-p**); quantification of tumour volumes in placebo ($n = 3$) (**h**) and pramlintide-treated mice ($n = 7$) (**q**), $P < 0.005$. Statistical significance indicated by black asterisk.



Extended Data Figure 8 | IAPP inhibits glycolysis by increasing intracellular G-6-P levels. **a, b**, Quantification of apoptosis (**a**) and proliferation (**b**), $n = 20$ fields of 3 biological replicates, $P < 0.005$. **c**, qRT-PCR for the target genes indicated on the x -axis in the indicated H1299 cells expressing the indicated siRNAs, $n = 4$. Asterisks indicate statistical significance ($P < 0.005$) relative to siNT. **d**, Western blot analysis of H1299 cells treated with the indicated siRNAs. **e, f**, Bar graph indicating glucose-dependent

proton secretion as a measure of glucose uptake and intracellular levels of glucose-6-phosphate in H1299 cells with the indicated siRNAs and treatments (**f**). **g**, Colour-coded legend for panels **e, f** and **i**. **h**, Western blot analysis of H1299 cells expressing the indicated siRNAs. **i**, Immunofluorescence analysis for ROS (red) or apoptosis (green or green/red) in H1299 cells expressing the indicated siRNAs and treated with 2DG and/or NAC.



Extended Data Figure 9 | Treatment of p53-mutant human cancer cell lines with pramlintide inhibits glycolysis and induces ROS and apoptosis.

a, b, Western blot analysis of H1299 cells expressing the indicated siRNAs (**a**) or concentrated media derived from H1299 cells expressing siNT, siΔNp63, or siΔNp73 (**b**). **c**, Extracellular acidification rate (ECAR) using H1299 cells expressing the indicated siRNAs and treated with the indicated media containing secreted IAPP and treated with the indicated amylin inhibitor (AI). **d–g**, Extracellular acidification rate (ECAR) as a measure of glycolysis in SW480 (**d**), MDA-MB-468 (**e**), SRB12 (**f**) and COLO16 (**g**) human cancer cell

lines after treatment with placebo, pramlintide, or pramlintide and a calcitonin receptor inhibitor (CalR I.), $n = 3$, $P < 0.005$. Glucose, oligomycin, and 2-deoxy-D-glucose (2DG) were supplied to the media at the indicated time points shown on the x-axis. **h, i**, Immunofluorescence for ROS (red) (**h**) and apoptosis (green) (**i**) on the indicated cells, $n = 3$. **j, k**, Kaplan–Meier survival curves using data from patients with p53 mutant tumours with the indicated cancers and co-expression of IAPP, RAMP3 and CALCR. Boxed numbers represent median survival.

The mitotic checkpoint complex binds a second CDC20 to inhibit active APC/C

Daisuke Izawa¹ & Jonathon Pines¹

The spindle assembly checkpoint (SAC) maintains genomic stability by delaying chromosome segregation until the last chromosome has attached to the mitotic spindle. The SAC prevents the anaphase promoting complex/cyclosome (APC/C) ubiquitin ligase from recognizing cyclin B and securin by catalysing the incorporation of the APC/C co-activator, CDC20, into a complex called the mitotic checkpoint complex (MCC). The SAC works through unattached kinetochores generating a diffusible 'wait anaphase' signal^{1,2} that inhibits the APC/C in the cytoplasm, but the nature of this signal remains a key unsolved problem. Moreover, the SAC and the APC/C are highly responsive to each other: the APC/C quickly targets cyclin B and securin once all the chromosomes attach in metaphase, but is rapidly inhibited should kinetochore attachment be perturbed^{3,4}. How this is achieved is also unknown. Here, we show that the MCC can inhibit a second CDC20 that has already bound and activated the APC/C. We show how the MCC inhibits active APC/C and that this is essential for the SAC. Moreover, this mechanism can prevent anaphase in the absence of kinetochore signalling. Thus, we propose that the diffusible 'wait anaphase' signal could be the MCC itself, and explain how reactivating the SAC can rapidly inhibit active APC/C.

The MCC is an APC/C inhibitor containing the MAD2, BUBR1 and BUB3 checkpoint proteins in a complex with CDC20⁵, where MAD2 and BUBR1 inhibit CDC20 by binding to substrate and APC/C recognition motifs^{6–8}. To elucidate how the SAC inhibits the APC/C we produced recombinant human MCC (rMCC) by co-expressing His₆-tagged MAD2, streptavidin binding protein (SBP)-tagged BUBR1 and untagged CDC20 at a 8:1:2 ratio (Extended Data Fig. 1a–e) in baculovirus-infected Sf9 cells. We co-purified MAD2, BUBR1 and CDC20 in a 'core MCC' complex at a 1:1:1 ratio (Extended Data Fig. 1b).

Incubating core rMCC with recombinant His₆-tagged CDC20 showed that core MCC could bind a second CDC20 molecule (Fig. 1a and Extended Data Fig. 1f), which was not due to CDC20 homodimerizing (Fig. 1a). Including BUB3 in the core rMCC made no difference to the amount of CDC20 that was bound (Extended Data Fig. 2). We note here recent speculation that the MCC may contain two molecules of CDC20⁹. The mode of binding to the second CDC20 differed from that required to form the core MCC because core MCC could bind to a CDC20^{AKILR} mutant unable to bind MAD2⁸ (Fig. 1a and Extended Data Fig. 1c). This also excluded the possibility that the second CDC20 had exchanged with CDC20 in the core MCC.

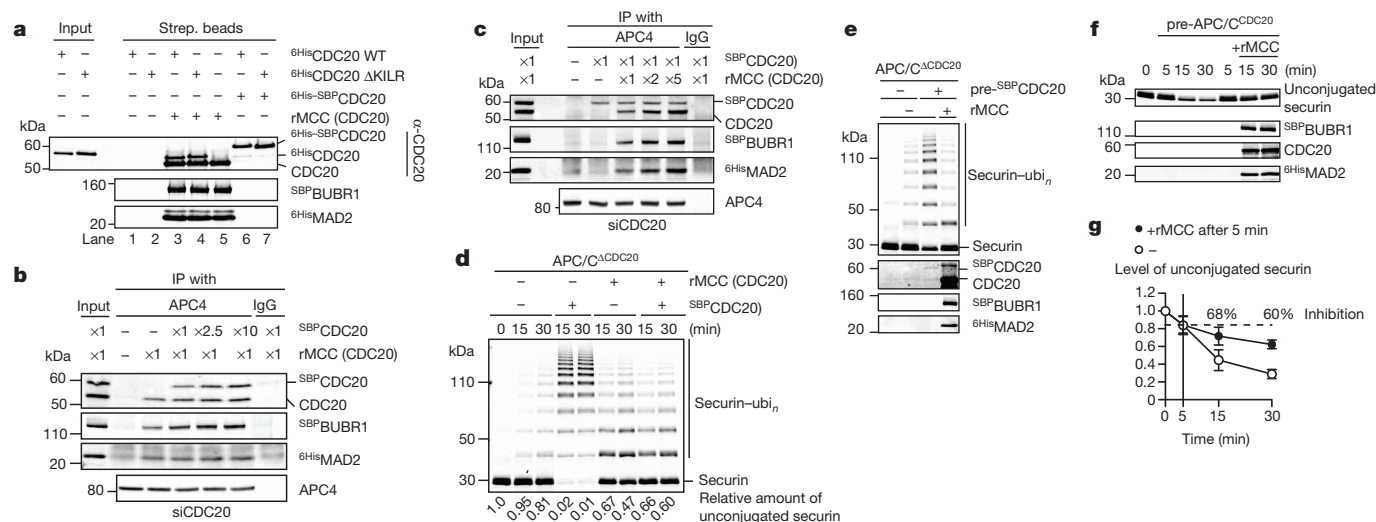


Figure 1 | Core MCC can inhibit APC/C^{CDC20}. **a**, Second CDC20 binding assay. ⁶His-SBP-CDC20 or rMCC, composed of untagged CDC20, SBP-BUBR1 and ⁶His-MAD2 were incubated with streptavidin (strep.) beads, unbound proteins washed away, and the beads incubated with either wild-type (WT) or ΔKILR (K¹²⁹ILR/AAAA) mutant ⁶His-CDC20 (Extended Data Fig. 1f). Proteins retained on the streptavidin beads were analysed by quantitative immunoblotting. Molecular mass markers are on the left; kDa, kilodalton. **b**, **c**, MCC prefers to bind APC/C^{CDC20}. The APC/C was immunoprecipitated from CDC20-depleted mitotic extracts supplemented with a constant amount of core MCC, and increasing amounts of ^{SBP}CDC20 (**b**), or vice versa (**c**), and analysed as in **a**. IP, immunoprecipitate. **d**, The MCC is an APC/C^{CDC20} inhibitor. The APC/C was immunoprecipitated as in **b** and incubated with infrared-dye-conjugated securin in an ubiquitylation reaction at 37 °C for 15 or

30 min with core rMCC and/or ^{SBP}CDC20 (1.5:1 ratio of core rMCC to rCDC20, see Extended Data Fig. 3a, b). Securin ubiquitylation (securin-ubi_n) was analysed by SDS-PAGE and a Li-COR Odyssey scanner. The amount of unconjugated securin is shown below the panel (level at 0 min is set to 1.0). **e–g**, The MCC inhibits active APC/C. **e**, The APC/C^{CDC20} was pre-incubated with ^{SBP}CDC20 to form APC/C^{CDC20}, unbound ^{SBP}CDC20 washed away, and APC/C^{CDC20} activity assayed as in panel **d** for 30 min. A 10-fold excess of rMCC to immunoprecipitated APC/C was added at 0 min (see also Extended Data Fig. 3c). **f**, APC/C activity was assayed as in **e** except that rMCC was added 5 min after starting the reaction. **g**, Unconjugated securin was measured from three independent experiments and the mean and s.d. plotted against time. To estimate APC/C inhibition, the level of securin at 5 min was set to 1.0. All results in Fig. 1 are representative of three or more experiments.

¹The Gurdon Institute and Department of Zoology, Tennis Court Road, Cambridge CB2 1QN, UK.

The question arose as to why we could not purify rMCC with two molecules of CDC20. We postulated that the second CDC20 bound less stably than the first CDC20, which is cooperatively bound by MAD2 and BUBR1⁶; therefore, limited amounts of CDC20 would preferentially incorporate into the core MCC. In agreement with this, we purified some core rMCC bound to a second CDC20 from Sf9 cell lysates containing excess CDC20 (50% bound in Extended Data Fig. 1g). We noted that increasing the amount of functional^{SBP} CDC20 enhanced core rMCC binding to the APC/C (Fig. 1b and Extended Data Fig. 1h, i). This indicated that core MCC could bind CDC20 associated with the APC/C, and that core rMCC did not compete with^{SBP} CDC20 for APC/C binding (Fig. 1c). This agreed with our previous finding that the MCC and CDC20 bind to the APC/C through different sites¹⁰.

To determine the properties of MCC as an APC/C^{CDC20} inhibitor we used a reconstituted ubiquitylation assay with APC/C isolated from CDC20-depleted mitotic cells (APC/C^{ΔCDC20}), and incubated it with^{SBP} CDC20 and/or core rMCC. Adding CDC20 strongly activated the APC/C, whereas, as expected^{6,8}, core MCC alone only weakly stimulated the APC/C (Fig. 1d). Neither MAD2 nor BUBR1 alone can inhibit the mitotic APC/C¹¹, and together they require pre-incubation to inhibit interphase APC/C^{CDC20} (ref. 7). By contrast, core MCC was a potent and rapid inhibitor of active APC/C^{CDC20}, as well as preventing CDC20 from activating the APC/C (Fig. 1d and Extended Data Fig. 3a, b), it inhibited active mitotic APC/C within 10 min (Fig. 1e–g and Extended Data Fig. 3c).

To gain insight into how core MCC could inhibit active APC/C^{CDC20}, we sought to identify how core MCC bound to a second CDC20. Studies on yeast MAD3/BUBR1 had implicated a number of D-box and KEN box motifs in binding to CDC20, and as important for the SAC^{12,13}. A D-box bound to the side of the CDC20 β-propeller domain in the MCC structure, whereas a KEN box bound to the top face⁶. We hypothesized that the second CDC20 might bind to the core MCC in a similar manner; therefore, we introduced mutations into the D-box receptor (D177A; ΔDR) and the KEN-box receptor (N329A/N331A/T377A/R445A; ΔKR) of CDC20. Both these CDC20 mutants bound much less well to core rMCC *in vitro* (Fig. 2a, b). Since the ΔDR mutant could still be incorporated into the core MCC (Fig. 2c), we tested whether inhibiting a second CDC20 was important for the SAC (Fig. 2d). We replaced endogenous CDC20 with the ΔDR mutant, or the ΔKR mutant as a positive control, and assayed the ability of cells to arrest in response to nocodazole. As expected, the ΔKR mutant abrogated the SAC because it could not form the core MCC (Fig. 2c–e). By contrast, the ΔDR mutant assembled into the core MCC and bound to the APC/C (Fig. 2c, d), yet the SAC was still defective (Fig. 2e). Cells expressing the ΔDR mutant, however, took more time to exit mitosis than those expressing the ΔKR mutant (Fig. 2e). We thought this might be because the ΔDR mutant was less effective at activating the APC/C¹⁴; consistent with this, cyclin B1 was degraded more slowly in these cells (Extended Data Fig. 4). These data supported the idea that the MCC inhibited a second CDC20 as part of a functional SAC.

Since CDC20 required its D-box and KEN box receptors to bind the core MCC, we identified the D-box and a KEN box on BUBR1 responsible for binding CDC20. The structure of the core MCC implicated a putative D-box⁶, but BUBR1 has two KEN-boxes: the first (K26EN) is essential to form the core MCC (Extended Data Fig. 1d), whereas the second (K304EN) is not required to form the core MCC but is still important for the SAC^{12,13,15,16}. We thought the second KEN box a more likely candidate to bind a second CDC20⁹; therefore, we mutated the putative D-box (R²²⁴xxL: ΔD-box) and the second KEN-box (ΔKEN2) in human BUBR1. Both mutants were incorporated into the core MCC *in vitro* (Fig. 3a and Extended Data Fig. 5a) and *in vivo* (Fig. 3c); both inhibited the CDC20 within the core MCC (Extended Data Fig. 5b), but reduced binding to a second molecule of CDC20 (Fig. 3a, b). (Note that BUBR1 alone did not bind two molecules of CDC20 because neither the D-box nor the second KEN-box was required to bind CDC20 in the absence of MAD2 (Extended Data Fig. 5c).) Furthermore, replacing endogenous

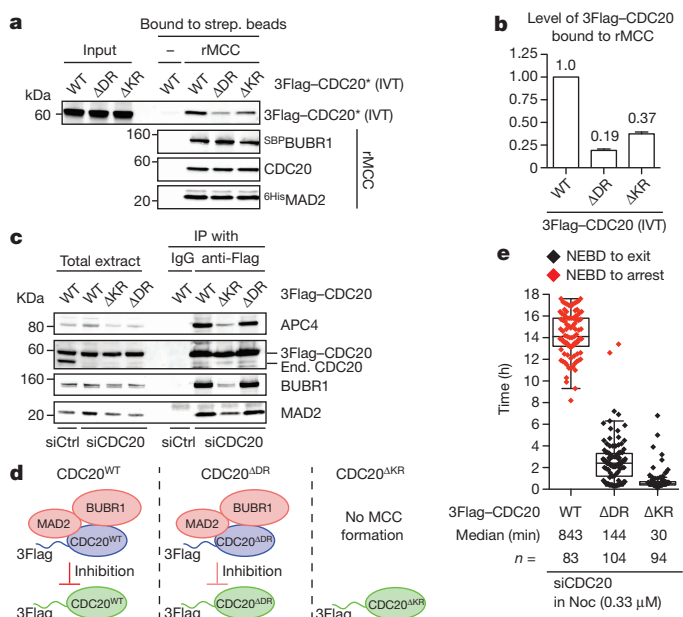


Figure 2 | The MCC binds to CDC20 through substrate recognition domains. **a**, Mutating CDC20 substrate recognition domains reduces binding to rMCC. **a**, rMCC was incubated with *in vitro* translated (IVT) 3×Flag-tagged wild-type CDC20, or CDC20^{ΔDR}, or CDC20^{ΔKR} mutants (indicated by *) and analysed as in Fig. 1a. ΔDR, D177 mutated to alanine; ΔKR, N329, N331, T377 and R445 mutated to alanine. **b**, Quantification of the data in panel **a** showing mean ± s.e.m. of three independent biological replicates. **c–e**, Defective SAC in cells expressing a CDC20 mutant that weakly binds the MCC. **c**, HeLa cells expressing siRNA-resistant 3×Flag-tagged CDC20 wild-type, or ΔKR, or ΔDR mutants, were treated with siRNA against CDC20, synchronized at prometaphase with nocodazole (Noc) and collected by mitotic shake-off. Anti-flag immunoprecipitates were analysed by quantitative immunoblotting. Results are representative of three biological replicates. Control cells (Ctrl) were depleted of GAPDH. End., endogenous. **d**, Schematic summary of CDC20^{ΔDR} and CDC20^{ΔKR} mutants. The CDC20^{ΔDR} mutant can form the MCC, but is only weakly bound and inhibited by the MCC. The CDC20^{ΔKR} mutant cannot form the MCC. **e**, HeLa cell lines of 3Flag-CDC20 were treated as in **c**, the time from nuclear envelope breakdown (NEBD) to mitotic exit measured, and plotted as a box and whisker chart where one diamond represents one cell. Red diamonds indicate the cell remained in mitosis until the end of the experiment. *n*, number of cells analysed in three independent experiments.

BUBR1 with the ΔD-box mutant (Fig. 3c) prevented cells from arresting in mitosis in response to either nocodazole (Fig. 3d), or Taxol where the SAC is much weaker¹⁷ (Extended Data Fig. 5d). Thus, the core MCC must inhibit a second CDC20 molecule to impose a functional SAC.

An important test of our idea that the core MCC inhibited active APC/C^{CDC20} was whether the core MCC could arrest a mitotic cell in which kinetochores could not catalyse further CDC20 incorporation into the core MCC (see Extended Data Fig. 6a). To prevent the core MCC from disassembling we attached a yellow fluorescent protein (YFP) tag to MAD2 (Venus–MAD2) and a green fluorescent protein (GFP)-binding domain (GBP)¹⁸ to CDC20 (GBP–CDC20). We called this stable complex MCC^{M2} (see Extended Data Figs 6 and 7); a similar approach using leucine zippers had been used previously in budding yeast¹⁹. We expressed MCC^{M2} in cells with normal levels of endogenous CDC20. MCC^{M2} was able to inhibit the APC/C when the SAC was inactivated in three different ways: (1) MCC^{M2} imposed a metaphase delay (Fig. 4a), in which the kinetochores did not stain for MAD2 (Extended Data Fig. 8a). The extent of the delay correlated with the amount of GBP–CDC20, and thus the amount of MCC^{M2} (Extended Data Fig. 8b); (2) MCC^{M2} imposed a delay in cells treated with the Mps1 kinase inhibitor reversine to prevent core MCC assembly²⁰ (Fig. 4b); (3) MCC^{M2} arrested cells in mitosis after depleting the KNL1 (also known as CASC5) kinetochore protein that

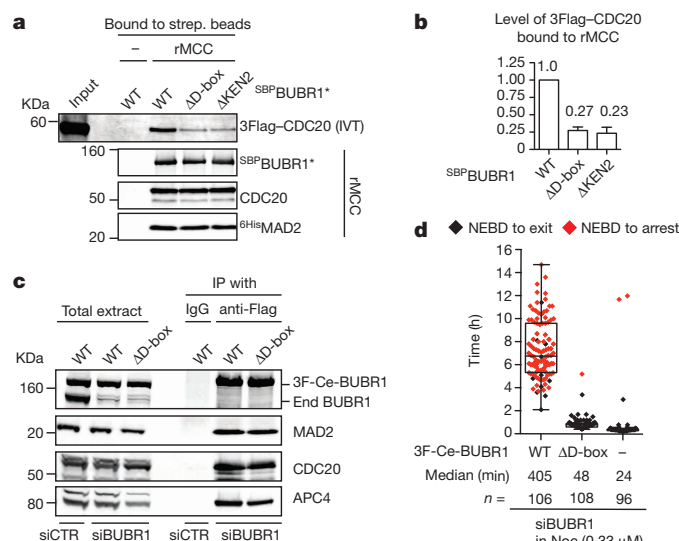


Figure 3 | MCC binds to the second CDC20 through the D-box of BUBR1 and this is required for the SAC. **a**, **b**, The D-box and second KEN box of BUBR1 bind to CDC20. **a**, rMCC containing^{SBP}BUBR1 wild-type, or Δ D-box, or Δ KEN2 (indicated by *), was incubated with 3 \times Flag-tagged CDC20 (IVT) and analysed as in **a**. Δ D-box, R224A, L226A; Δ KEN2, K304EN mutated to AAA. **b**, Quantification of the data in panel **a** to show the mean \pm s.e.m. of four independent biological replicates. **c**, The D-box mutant of BUBR1 forms the MCC. HeLa cells expressing siRNA-resistant 3 \times Flag-Cerulean-BUBR1 (3F-Ce-BUBR1), either wild-type or the Δ D-box mutant, were treated with siRNA against BUBR1, and prometaphase cells collected by mitotic shake off and analysed as in Fig. 2c. Result is representative of three biological replicates. **d**, The D-box of BUBR1 is required for the SAC. HeLa cell lines expressing wild-type and Δ D-box mutant 3F-Ce-BubR1 were treated as in **c**, 0.33 μ M nocodazole was added and the time from NEBD to mitotic exit was measured as in Fig. 2e. **n**, number of analysed cells from two independent biological replicates. Note that the Δ D-box mutation did not affect the recruitment of BUBR1 to unattached kinetochores (Extended Data Fig. 5e).

is required for the SAC²¹ (Extended Data Fig. 8c–e). These data supported our idea that the core MCC inhibited active APC/C^{CDC20}. Moreover, as the MCC inhibits the APC/C without further signalling from the kinetochores, it has one of the essential properties required of the diffusible ‘wait anaphase’ inhibitor, although our data do not prove that it is the diffusible inhibitor *in vivo*.

All the functional components of the core MCC were required for MCC^{M2} to inhibit APC/C^{CDC20} because we could not delay cells in mitosis when we stabilized the binding between MAD2 and CDC20 in the absence of BUBR1 (Fig. 4c), nor when we stabilized MAD2 with a CDC20– Δ KILR mutant that cannot form the core MCC (Extended Data Fig. 9a). Finally, we stabilized the binding between MAD2 and CDC20 (MCC^{M2}), but replaced BUBR1 with the Δ D-box mutant to perturb binding to a second CDC20. These complexes were much less effective at inhibiting APC/C^{CDC20} *in vitro* (Extended Data Fig. 9b), and unable to delay cells in mitosis (Fig. 4c; model in Extended Data Fig. 9c). Thus, we conclude that to arrest cells in mitosis the core MCC inhibits a second molecule of CDC20 that can even be part of an active APC/C^{CDC20}.

Crucial gaps have remained in our understanding of the SAC: notably, how the ‘wait anaphase’ signal generated at unattached kinetochores inhibits APC/C activity in the rest of the cell². Unattached kinetochores appear to catalyse a conformational change in MAD2 to bind CDC20²² and subsequently promote APC/C–MCC formation in the cytoplasm. However, it is unlikely that all CDC20 could be bound by MAD2 at the kinetochore, therefore additional mechanisms have been proposed to prevent the activation of the APC/C, including cytoplasmic amplification of MAD2–CDC20 binding²², although this now appears unlikely^{19,23}, and phosphorylation of CDC20 by BUB1²⁴. We now show how the MCC, formed either at kinetochores or in the cytoplasm, could act as a diffusible

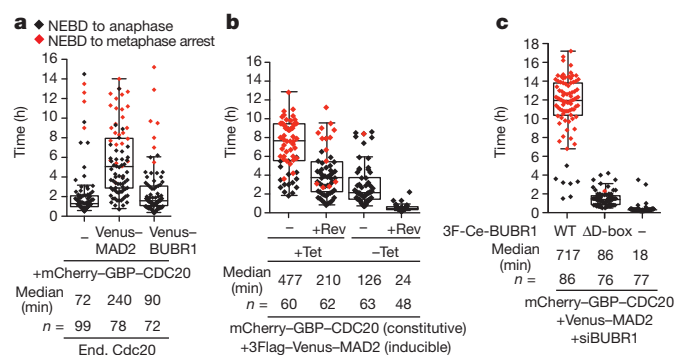


Figure 4 | A stabilized MCC delays anaphase by inhibiting endogenous APC/C^{CDC20}. **a**, HeLa cells were transfected with plasmids encoding mCherry–GBP–CDC20 and either Venus–MAD2 or Venus–BUBR1, and the time from NEBD to anaphase was analysed in unperturbed mitoses as in Fig. 2e. **n**, number of cells from three independent biological replicates. **b**, HeLa cell lines stably expressing mCherry–GBP–CDC20 and a tetracyclin-inducible 3 \times Flag–Venus–MAD2 were treated, or not, with 1 μ M reversine (+Rev) 6 h after release from a thymidine block in the presence (+Tet) or absence (–Tet) of tetracyclin, and analysed as in **a**. **n**, number of cells from two independent biological replicates. **c**, HeLa cells lines in Fig. 3c expressing 3F-Ce-BUBR1 plus mCherry–GBP–CDC20 and Venus–MAD2 were treated with siRNA against BUBR1 and analysed as in panel **a**. **n**, number of cells from two independent biological replicates.

inhibitor to inhibit APC/C^{CDC20} throughout the cell (Extended Data Fig. 10), although our data do not prove that it disseminates the ‘wait anaphase’ signal *in vivo*. Previously, it has been proposed⁷ that the complex between MAD2 and CDC20 will template the formation of the BBC (BUBR1–BUB3–CDC20) complex²⁵ to inhibit CDC20—although in these experiments p31^{Comet} was depleted, which would alter the levels and behaviour of checkpoint complexes^{25–27}. While we also find that the BBC is an abundant APC/C inhibitor in cells^{27,28}, we show here that stabilizing the MCC generates a more potent inhibitor than stabilizing the BBC (Fig. 4a; MCC^{R1} see Extended Data Fig. 6b), which agrees with the observation that cells containing a greater proportion of MCC over BBC exhibit stronger SAC activity^{17,25}. Our results could also solve a further conundrum posed by the SAC. MAD2 and the APC/C bind to the same KILR motif on CDC20⁸; therefore, CDC20 must dissociate from the APC/C to bind MAD2. By analogy with measurements on Cdh1²⁹, CDC20 is predicted to dissociate slowly from the APC/C (half time of dissociation \sim 25 min), yet reactivating the SAC can inhibit active APC/C in less than 5 min^{3,4}. Our finding that MCC rapidly inhibits CDC20 already bound to the APC/C can help to explain the close temporal coupling between the SAC and the APC/C. Indeed, our data indicate that the MCC prefers CDC20 that is already bound to the APC/C; the reason for this will be important to determine in the future.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 21 February; accepted 2 October 2014.

Published online 12 November 2014.

- Rieder, C. L., Cole, R. W., Khodjakov, A. & Sluder, G. The checkpoint delaying anaphase in response to chromosome monoorientation is mediated by an inhibitory signal produced by unattached kinetochores. *J. Cell Biol.* **130**, 941–948 (1995).
- Rieder, C. L. *et al.* Mitosis in vertebrate somatic cells with two spindles: implications for the metaphase/anaphase transition checkpoint and cleavage. *Proc. Natl Acad. Sci. USA* **94**, 5107–5112 (1997).
- Clute, P. & Pines, J. Temporal and spatial control of cyclin B1 destruction in metaphase. *Nature Cell Biol.* **1**, 82–87 (1999).
- Dick, A. E. & Gerlich, D. W. Kinetic framework of spindle assembly checkpoint signalling. *Nature Cell Biol.* **15**, 1370–1377 (2013).
- Sudakin, V., Chan, G. K. & Yen, T. J. Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2. *J. Cell Biol.* **154**, 925–936 (2001).

6. Chao, W. C., Kulkarni, K., Zhang, Z., Kong, E. H. & Barford, D. Structure of the mitotic checkpoint complex. *Nature* **484**, 208–213 (2012).
7. Han, J. S. *et al.* Catalytic assembly of the mitotic checkpoint inhibitor BubR1-Cdc20 by a Mad2-induced functional switch in Cdc20. *Mol. Cell* **51**, 92–104 (2013).
8. Izawa, D. & Pines, J. Mad2 and the APC/C compete for the same site on Cdc20 to ensure proper chromosome segregation. *J. Cell Biol.* **199**, 27–37 (2012).
9. Primorac, I. & Musacchio, A. Panta rhei: the APC/C at steady state. *J. Cell Biol.* **201**, 177–189 (2013).
10. Izawa, D. & Pines, J. How APC/C-Cdc20 changes its substrate specificity in mitosis. *Nature Cell Biol.* **13**, 223–233 (2011).
11. Tang, Z., Bharadwaj, R., Li, B. & Yu, H. Mad2-independent inhibition of APC/Cdc20 by the mitotic checkpoint protein BubR1. *Dev. Cell* **1**, 227–237 (2001).
12. Burton, J. L. & Solomon, M. J. Mad3p, a pseudosubstrate inhibitor of APC/Cdc20 in the spindle assembly checkpoint. *Genes Dev.* **21**, 655–667 (2007).
13. King, E. M., van der Sar, S. J. & Hardwick, K. G. Mad3 KEN boxes mediate both Cdc20 and Mad3 turnover, and are critical for the spindle checkpoint. *PLoS ONE* **2**, e342 (2007).
14. Tian, W. *et al.* Structural analysis of human Cdc20 supports multisite degron recognition by APC/C. *Proc. Natl Acad. Sci. USA* **109**, 18419–18424 (2012).
15. Elowe, S. *et al.* Uncoupling of the spindle-checkpoint and chromosome-congression functions of BubR1. *J. Cell Sci.* **123**, 84–94 (2010).
16. Lara-Gonzalez, P., Scott, M. I., Diez, M., Sen, O. & Taylor, S. S. BubR1 blocks substrate recruitment to the APC/C in a KEN-box-dependent manner. *J. Cell Sci.* **124**, 4332–4345 (2011).
17. Collin, P., Nashchekina, O., Walker, R. & Pines, J. The spindle assembly checkpoint works like a rheostat rather than a toggle switch. *Nature Cell Biol.* **15**, 1378–1385 (2013).
18. Rothbauer, U. *et al.* A versatile nanotrapp for biochemical and functional studies with fluorescent fusion proteins. *Mol. Cell. Proteomics* **7**, 282–289 (2008).
19. Lau, D. T. & Murray, A. W. Mad2 and Mad3 cooperate to arrest budding yeast in mitosis. *Curr. Biol.* **22**, 180–190 (2012).
20. Santaguida, S., Tighe, A., D'Alise, A. M., Taylor, S. S. & Musacchio, A. Dissecting the role of MPS1 in chromosome biorientation and the spindle checkpoint through the small molecule inhibitor reversine. *J. Cell Biol.* **190**, 73–87 (2010).
21. Kiyomitsu, T., Obuse, C. & Yanagida, M. Human blinkin/AF15q14 is required for chromosome alignment and the mitotic checkpoint through direct interaction with Bub1 and BubR1. *Dev. Cell* **13**, 663–676 (2007).
22. De Antoni, A. *et al.* The Mad1/Mad2 complex as a template for Mad2 activation in the spindle assembly checkpoint. *Curr. Biol.* **15**, 214–225 (2005).
23. Mariani, L. *et al.* Role of the Mad2 dimerization interface in the spindle assembly checkpoint independent of kinetochores. *Curr. Biol.* (2012).
24. Tang, Z., Shu, H., Oncel, D., Chen, S. & Yu, H. Phosphorylation of Cdc20 by Bub1 provides a catalytic mechanism for APC/C inhibition by the spindle checkpoint. *Mol. Cell* **16**, 387–397 (2004).
25. Westhorpe, F. G., Tighe, A., Lara-Gonzalez, P. & Taylor, S. S. p31 comet-mediated extraction of Mad2 from the MCC promotes efficient mitotic exit. *J. Cell Sci.* **124**, 3905–3916 (2011).
26. Varet, G., Guida, C., Santaguida, S., Chiroli, E. & Musacchio, A. Homeostatic control of mitotic arrest. *Mol. Cell* **44**, 710–720 (2011).
27. Mansfeld, J., Collin, P., Collins, M. O., Choudhary, J. & Pines, J. APC15 drives the turnover of MCC-Cdc20 to make the spindle assembly checkpoint responsive to kinetochore attachment. *Nature Cell Biol.* **13**, 1234–1243 (2011).
28. Nilsson, J., Yekezare, M., Minshull, J. & Pines, J. The APC/C maintains the spindle assembly checkpoint by targeting Cdc20 for destruction. *Nature Cell Biol.* **10**, 1411–1420 (2008).
29. Matyskiela, M. E. & Morgan, D. O. Analysis of activator-binding sites on the APC/C supports a cooperative substrate-binding mechanism. *Mol. Cell* **34**, 68–80 (2009).

Acknowledgements We are grateful to T. Matsusaka for developing infrared-dye-conjugated ubiquitylation substrates, to A. Musacchio, W. Earnshaw, T. Kiyomitsu and M. Yanagida for reagents, and to A. Musacchio and members of our laboratory for critical discussions. This work was supported by a project and a programme grant from Cancer Research UK to J.P. J.P. acknowledges core funding to the Gurdon Institute from the Wellcome Trust and CR UK.

Author Contributions Experiments were designed by D.I. and J.P., carried out by D.I., and analysed by D.I. and J.P.; D.I. and J.P. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.P. (jp103@cam.ac.uk).

METHODS

Cell culture and synchronization. HeLa cells were maintained in Advanced D-MEM with 2% FBS. For synchronisation at the beginning of S phase, HeLa cells were treated with 2.5 mM thymidine as previously described⁸. For prometaphase, cells were released from a thymidine block and 6 h later treated with nocodazole at a final concentration of 0.33 μ M for 6–12 h. For SAC-inactivated samples, cells were released from a nocodazole block into medium including 1 μ M reversine and 10 μ M MG132 for a further 1 h.

Transfection with siRNA and DNA. The following ON-TARGETplus (Dharmacon, CO, USA) oligonucleotides as previously described⁸ were used: *CDC20* 50 nM (CGGAAGACCUGCCGUUACAUAU); *MAD2* 20 nM (GGAAGAGUCGGGACCACAGUU); *BUBR1* 50 nM (GAUGGUGAAUUGUGGAAUA); *KNL1* 50 nM (AAGAUCUGAUUAAGGAUCCACGAAA) and *GAPDH* (D-001830-01). Cells were transfected with short interfering RNA (siRNA) oligonucleotides once or twice at the indicated concentrations using lipofectamine RNAiMax (Invitrogen). To transfect siRNA oligonucleotides and DNA plasmids at the same time, cells were treated with lipofectamine 2000 (Invitrogen). An siRNA-resistant of resistant open reading frame (ORF) of *BUBR1* is generated by mutating underlined nucleotides (GATGGCGAGCTTUGGAAUA).

In vitro reconstituted ubiquitylation assay. *In vitro* ubiquitylation assays were performed as described previously⁸ but with modifications to use a fluorescently-labelled substrate developed by T. Matsusaka. In brief, CDC20 was depleted by siRNA treatment for 48 h before the APC/C was purified with anti-APC3 (AF3.1) antibody from mitotic HeLa cell extract. Immunoprecipitates were resuspended in ubiquitylation reaction buffer contained E1-ligase, UbcH10 (E2), ubiquitin, ATP, ATP regenerating system, and fluorescently-labelled securin as a substrate in QA buffer (100 mM NaCl, 30 mM Hepes pH 7.8, 2 mM ATP, 2 mM MgCl₂, 0.1 μ g μ l⁻¹ BSA, 1 mM DTT) at 37 °C for the indicated time, and supplied with recombinant CDC20 and/or core rMCC as indicated. Recombinant securin protein was labelled with IRDye680 dye (IRDye 680LT Maleimide Infrared Dye; LiCOR) according to the manufacturer's instructions and directly scanned with a Li-COR Odyssey CCD scanner after SDS-PAGE analysis. Ubiquitylation of CDC20, MAD2 and BUBR1 were analysed by quantitative immunoblotting. After blotting with primary antibodies, blots were incubated with fluorescently labelled secondary antibodies and the fluorescence measured using a LI-COR Odyssey CCD scanner according to the manufacturer's instructions (LI-COR Biosciences, NE, USA).

Expression of mCherry-GBP-CDC20, Venus-BUBR1 and Venus-MAD2. We used two types of human expression vectors: pcDNA5-3Flag-Venus (inducible CMV promoter) and pmCherry-CAG-C1 (chicken β -actin promoter). In the pcDNA5-3Flag-Venus, 3Flag-Venus is inserted into the multiple-cloning site of pcDNA5/FRT/TO (Invitrogen). In the pmCherry-CAG-C1 vector, EYFP and CMV promoter of pEYFP-C1 (Clontech) were replaced by mCherry and CAG promoter, respectively. A siRNA-resistant open reading frame (ORF) of *CDC20* and *GBP-CDC20* were cloned into the pmCherry-CAG-C1 vector, and *MAD2* and *BUBR1* were cloned into pcDNA5-3Flag-Venus. All constructs were verified by sequencing and sequences are available on request. To co-express mCherry-GBP-CDC20 and Venus-MAD2 or Venus-BUBR1, the indicated plasmids were co-transfected with the indicated siRNA oligonucleotides using Lipolipofectamine 2000 (Invitrogen).

Inducible cell lines. To generate cell lines expressing 3Flag-CDC20 or 3Flag-Cerulean-BUBR1 proteins from an inducible promoter, a siRNA-resistant ORF of *CDC20* or *BUBR1* was cloned into a modified version of pcDNA5/FRT/TO (Invitrogen). Those plasmids were transfected into a HeLa-FRT cell line (a gift from S. Taylor) and stable cell lines were generated using the FLIP-in system (Invitrogen). To obtain a cell line expressing 3Flag-Venus-MAD2 from an inducible CMV promoter and mCherry-GBP-CDC20 from a constitutive CMV promoter (used in Fig. 4b), a HeLa-FRT cell line expressing an inducible 3Flag-Venus-MAD2 was

transfected with the pmCherry-C1-GBP-CDC20 plasmid and selected with Geneticin (Invitrogen). To induce proteins from the inducible promoter, cells were treated with tetracycline (1 μ g ml⁻¹, Calbiochem) 36 h before analysis.

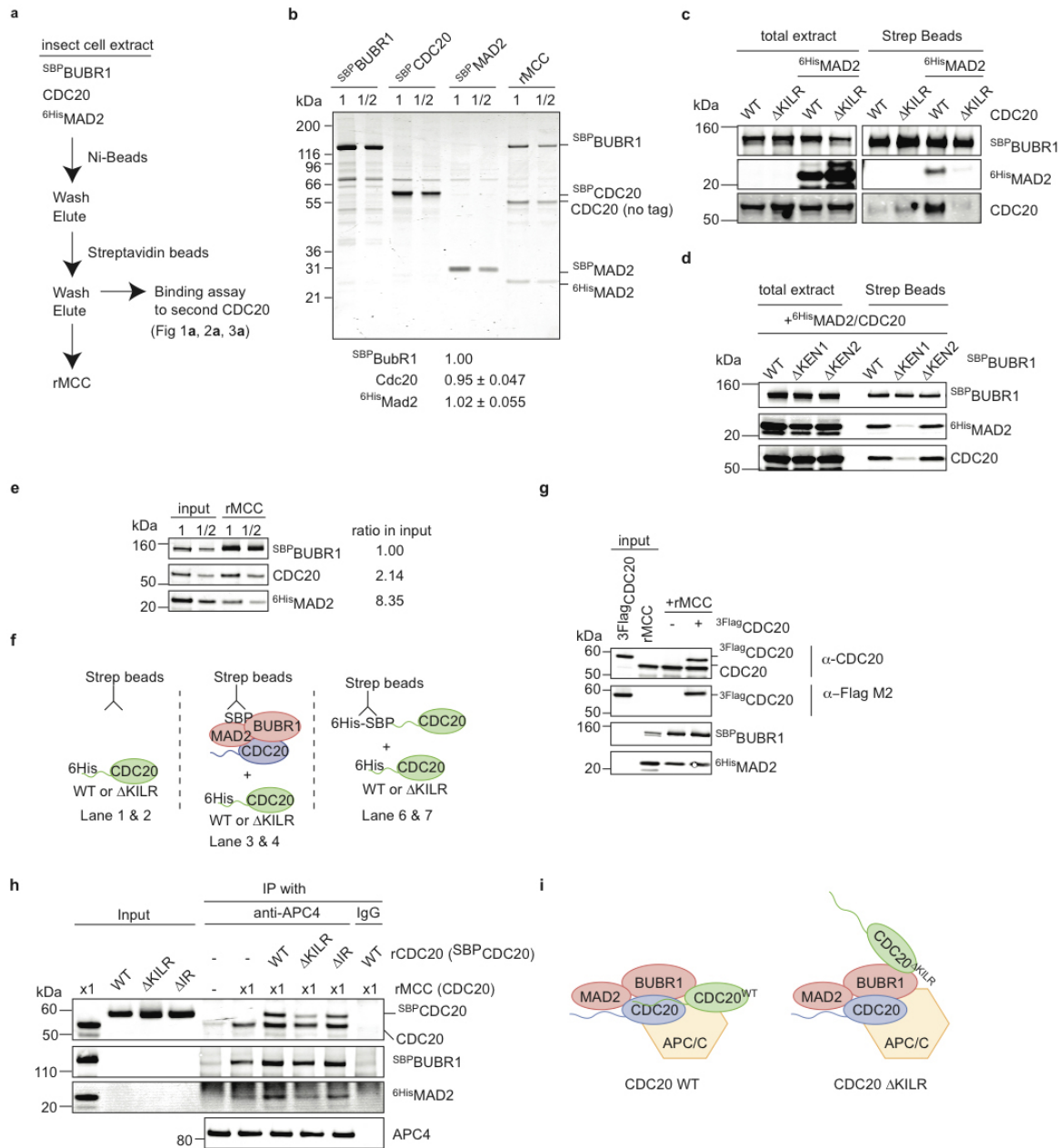
Immunoprecipitation and size exclusion chromatography. Cells for immunoprecipitation were lysed with HEPES buffer (150 mM KCl, 20 mM Hepes pH 7.8, 10 mM EDTA, 10% glycerol, 0.2% NP-40, 1 mM dithiothreitol (DTT), Roche complete inhibitor cocktail tablet, 0.2 μ M microcystin, 1 mM PMSF) for 10 min on ice and clarified by a 20,000g spin for 10 min. Protein complexes were immunoprecipitated with antibodies (anti-APC4, anti-APC3 (AF3.1), anti-GFP or anti-flag M2 epitope) covalently coupled to Protein G Dynabeads (Invitrogen) using HEPES buffer for incubation and washing. For size exclusion chromatography analysis, cells were pelleted then resuspended in buffer A (140 mM NaCl, 20 mM Hepes pH 7.8, 6 mM MgCl₂, 5% glycerol, 1 mM DTT, Roche complete inhibitor cocktail tablet, 0.2 μ M microcystin, 1 mM PMSF) at a 1:1 ratio of buffer to cells, and lysed by nitrogen cavitation (1,000 p.s.i., 30 min, Parr Instruments, USA). Lysed cells were centrifuged at 20,000g for 10 min and 259,000g for 10 min before loading onto a Superose 6 PC 3.2/30 column (GE Healthcare). The column was run at a flow rate of 25 μ l min⁻¹ in buffer B (140 mM NaCl, 30 mM Hepes pH 7.8, 5% glycerol, 1 mM DTT) and 50- μ l fractions collected.

Epifluorescence. Cells were seeded into 8-well dishes (Thistle Scientific, UK) to enable experiments to be performed in parallel. Before imaging, the culture medium was replaced with Leibovitz's L-15 medium (Gibco Life Technologies, UK) supplemented with 10% fetal bovine serum and penicillin/streptomycin. Cells were imaged on a DeltaVision microscope equipped with an environmental chamber at 37 °C (API, USA) with a QuantEM camera (Photometrics, USA) and Lambda LS illumination (Sutter, USA) as previously described¹⁰, or a spinning disc microscope (Intelligent Imaging Innovations, Colorado, USA) equipped with a CSU-X1 head (Yokogawa, Japan) and a QuantEM:512sc EMCCD camera (Photometrics, USA). In Figs 2e, 3d and 4, images of DIC and fluorescence were captured at 6-min intervals and the fluorescence intensities were measured and analysed using ImageJ/Fiji software as previously described¹⁰.

Antibodies. The following antibodies were used at the indicated dilutions. CDC20 (sc-13162, Santa Cruz Biotechnology) 1:500; CDC20 (A301-180A, Bethyl laboratories) 1:500; BUBR1 (612503, BD transduction laboratories); BUBR1 (A300-386A, Bethyl laboratories) 1:500; MAD2 (610679, BD transduction laboratories) 1:500; MAD2 (A300-301A, Bethyl Laboratories) 1:500; BUB3 (611730, BD Transduction Laboratories) 1:500; APC3 (610455, BD Transduction Laboratories) 1:500; APC4 (monoclonal antibody raised against a carboxy-terminal peptide) 1:500; KNL1 (a gift from M. Yanagida and T. Kiyomitsu) 1:50; anti-myc-epitope (9E10, Santa Cruz Biotechnology) 1:500; anti-flag epitope (M2, Sigma) 1:5,000; anti-GFP (Clone 3.1 and 7.1, Roche) 1:200.

Secondary antibodies: IRDye 680CW donkey anti-mouse (926-68072, LI-COR), IRDye 800CW donkey anti-mouse (926-32212, LI-COR); IRDye 680CW donkey anti-rabbit (926-32223, LI-COR); IRDye 800CW donkey anti-rabbit (926-32213, LI-COR) were all used at 1:10,000.

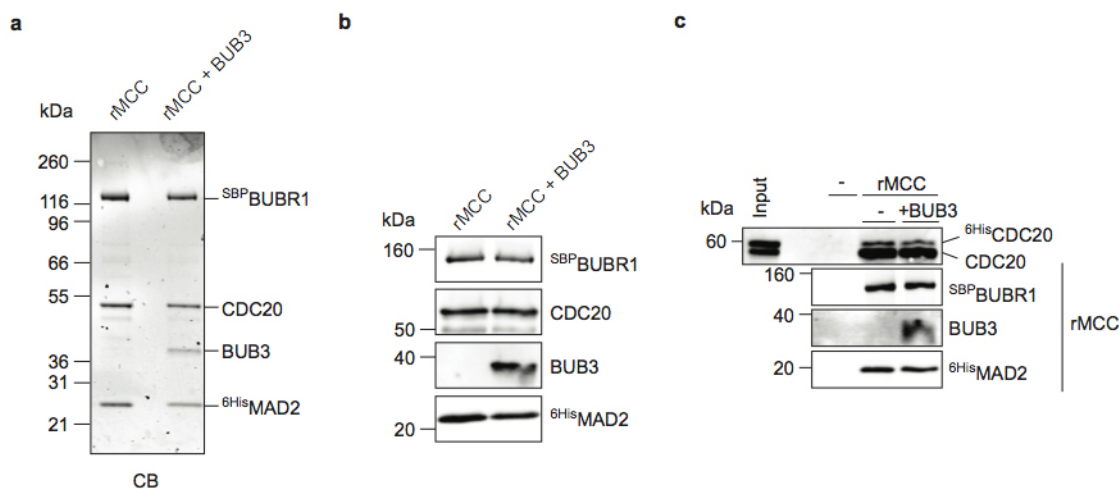
Immunofluorescence. Cells were fixed in 4% paraformaldehyde and 2% sucrose for 5 min. After fixation cells were blocked in 3% BSA-PBS NP-40 0.2% and then incubated with antibodies. All antibodies were diluted in 3% BSA-PBS NP-40 0.2% and washes were performed with PBS NP-40 0.2%. Antibodies were used at the following dilutions: anti-flag M2 (sigma), 1:4000; anti-GFP (Roche) and anti-CDC20 (sc-13162, Santa Cruz Biotechnology), 1:400; anti-MAD2 (A300-301A, Bethyl Laboratories) and anti-BUBR1 (A300-301A, Bethyl Laboratories), 1:200. Anti-ACA serum (a gift from W. Earnshaw) was used at 1:20,000. Secondary antibodies conjugated to Alexa Fluor 488, Alexa Fluor 568 or Alexa Fluor 647 (Molecular Probes) were diluted 1:400. DNA was stained with Hoechst-33342.



Extended Data Figure 1 | Recombinant human mitotic checkpoint complex binds to a second CDC20.

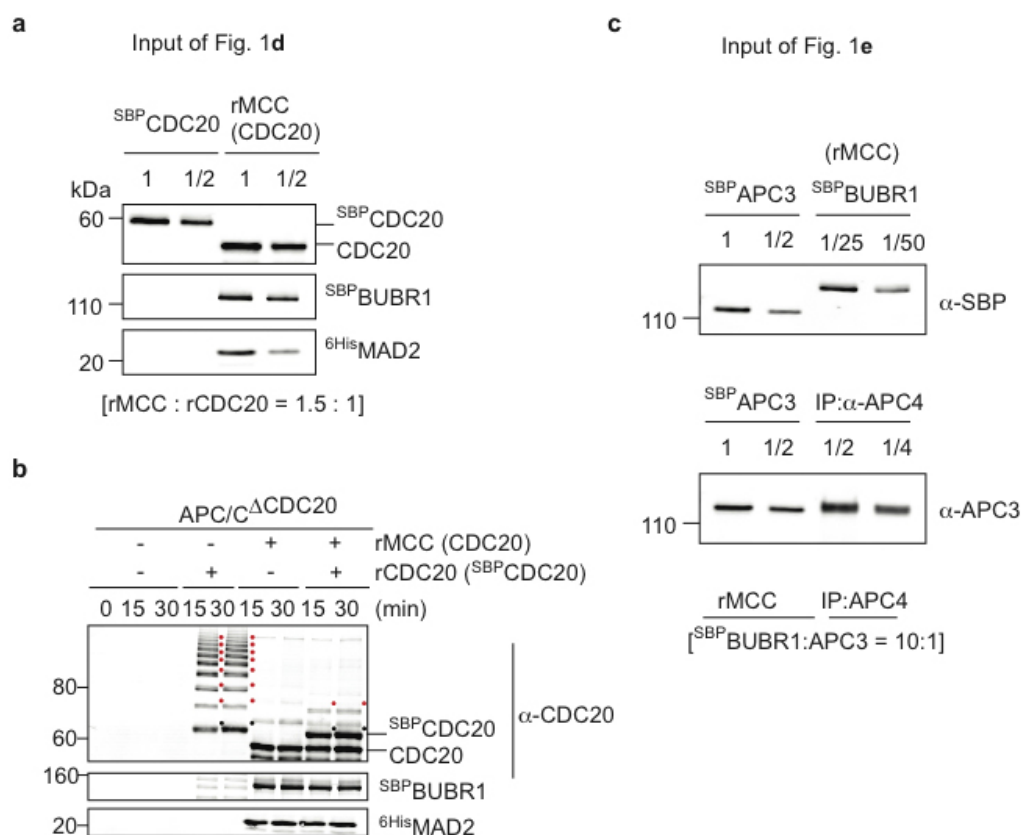
a, Schematic illustration of purification steps. Human wild-type CDC20 (untagged), ^{SBP}BUBR1 and ^{6His}MAD2 were expressed in baculovirus-infected Sf9 cells. The recombinant core mitotic checkpoint complex (rMCC) was purified by nitrilotriacetic acid (Ni-NTA) and streptavidin beads. Purified core rMCC bound to streptavidin beads was used to assay binding to purified recombinant Cdc20. **b**, Core rMCC consisting of CDC20, ^{SBP}BUBR1 and ^{6His}MAD2 was analysed by SDS-PAGE and Coomassie blue R250 staining, followed by quantification at 680 nm on a LiCOR Odyssey scanner. Equal molar amounts of purified ^{SBP}BUBR1, ^{SBP}CDC20 and ^{SBP}MAD2 proteins were used to calibrate the Coomassie blue staining. The stoichiometry of core rMCC (mean ± s.d. is shown below the panel with ^{SBP}BUBR1 set to 1.0) was estimated from three independently purified core rMCC preparations. Molecular mass markers are on the left. **c, d**, Both the MAD2 binding motif of CDC20 and the first KEN box of BUBR1 are required to assemble rMCC. Core rMCC was pulled down with streptavidin beads from Sf9 cells expressing ^{SBP}BUBR1, ^{6His}MAD2 and either wild-type (WT) CDC20 or the K129ILR/AAAA mutant (ΔKILR) (**c**), or ^{6His}MAD2, wild-type CDC20 plus wild-type ^{SBP}BUBR1, or alanine substitution mutants of either KEN box 1 (ΔKEN1) or KEN box 2 (ΔKEN2). The proteins retained on streptavidin beads were analysed by immunoblotting with the indicated antibodies. **e**, Relative expression levels of core rMCC components. Sf9 cells

extracts expressing the core rMCC, and the purified rMCC complex, were analysed by quantitative immunoblotting. The ratio of the proteins in the extracts is given, with that of ^{SBP}BUBR1 set to 1.0. **f**, Schematic illustration of the second CDC20 binding assay in Fig. 1a. In lanes 1 and 2, the streptavidin beads were incubated with either ^{6His}CDC20 wild-type or the ΔKILR (K129ILR/AAAA) mutant. In lanes 3 and 4, the streptavidin beads bound to core rMCC were incubated with the ^{6His}CDC20 proteins. In lanes 6 and 7, the streptavidin beads bound to ^{6His}-SBP-CDC20 were incubated with the ^{6His}CDC20 proteins. **g**, Sf9 cell extracts expressing core rMCC or 3Flag-tagged CDC20 were mixed and the core rMCC purified as in **a**. The core rMCC was analysed by quantitative immunoblotting. 51% of the core rMCC was purified bound to a second ^{3Flag}CDC20. **h**, A functional CDC20 promotes the binding of core rMCC to the APC/C. The APC/C was immunoprecipitated from CDC20-depleted mitotic extracts supplemented with a constant amount of core rMCC and tenfold excess of recombinant wild-type ^{SBP}CDC20, or the ΔKILR or ΔIR mutants. The co-immunoprecipitates were analysed as in Fig. 1c. **i**, Schematic of the APC/C-MCC-CDC20 ternary complex. Both core rMCC and CDC20 bind to the APC/C and form a ternary complex (left). The CDC20^{ΔKILR} mutant cannot bind the APC/C directly, nor stimulate core rMCC binding to the APC/C, but CDC20^{ΔKILR} still binds to rMCC (right). All results are representative of two or more independent biological replicates.



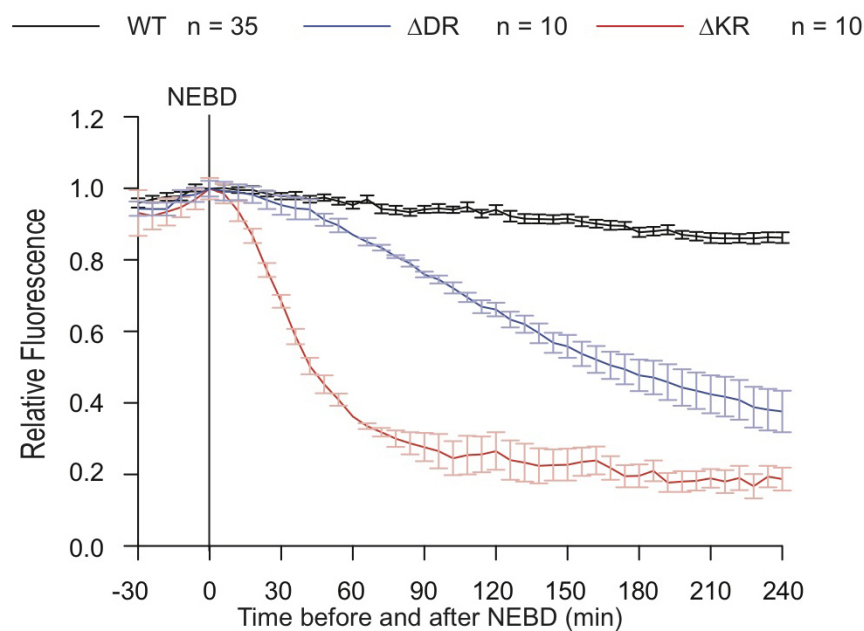
Extended Data Figure 2 | Comparison of rMCC with and without BUB3.
a, b, Preparation of recombinant core MCC with or without BUB3. Insect cells were infected with viruses expressing core MCC components with and without BUB3, and the rMCC was purified by Ni-NTA and streptavidin beads. The

complexes were analysed by Coomassie blue (CB) staining (**a**) and immunoblotting (**b**). **c,** Binding to a second ⁶His-CDC20 of recombinant core MCC with or without BUB3 was performed and analysed as in Fig. 1a. All results are representative of two independent biological replicates.



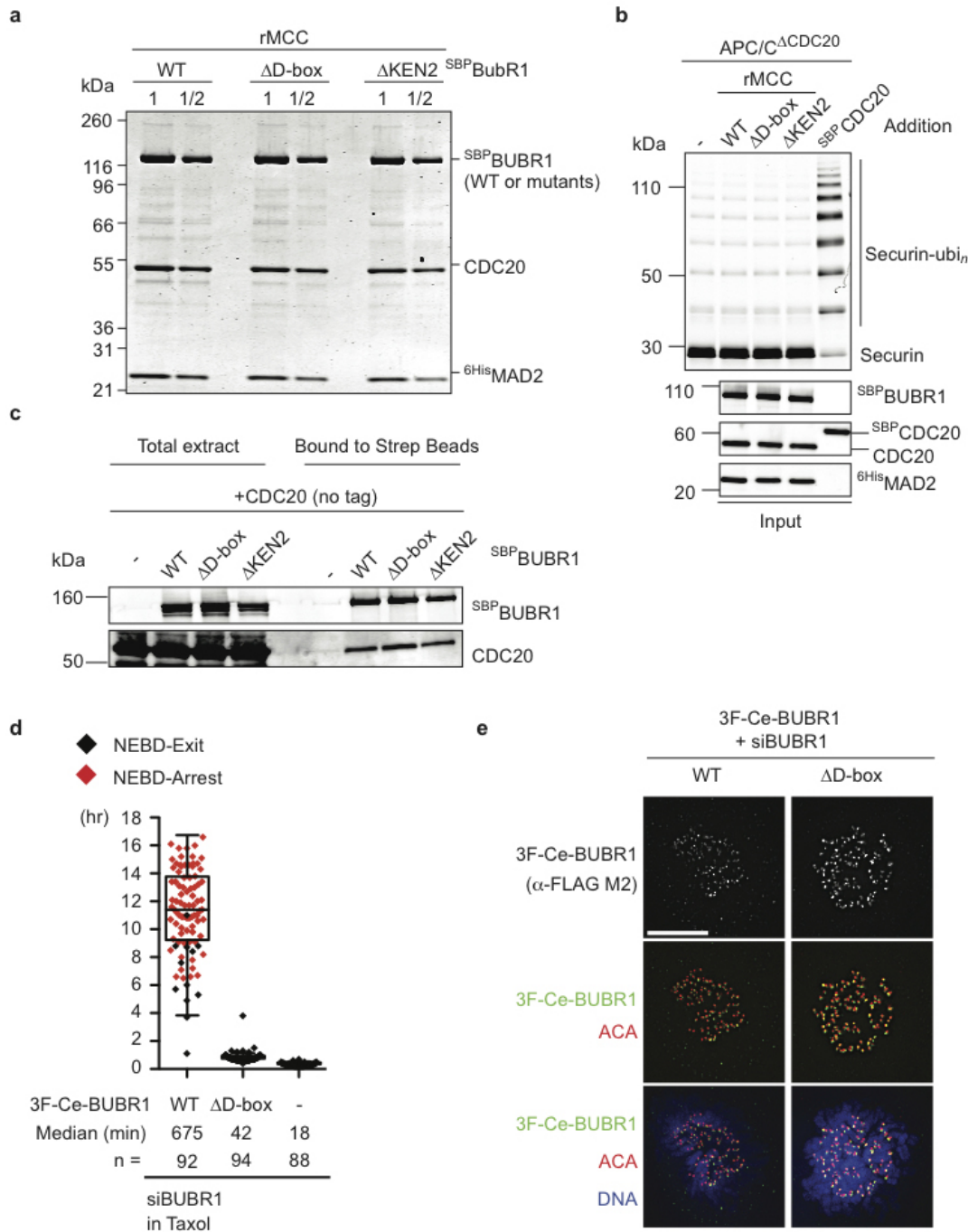
Extended Data Figure 3 | Molar ratios of rMCC, CDC20 and the APC/C in the *in vitro* ubiquitylation assays. **a, b,** Core rMCC and CDC20 from Fig. 1d were analysed by quantitative immunoblotting. CDC20, MAD2 and BUBR1 were analysed by quantitative immunoblotting in the input (**a**) and in the reaction (**b**). The black filled circles are unconjugated ^{SBP}CDC20; red filled

circles are ubiquitylated ^{SBP}CDC20. **c,** Core rMCC, rCDC20, and the APC/C immunoprecipitates used in Fig. 1e, plus a purified ^{SBP}APC3 subunit, were analysed by quantitative immunoblotting with the indicated antibodies. The calculated molar ratios of rMCC, rCDC20 and the APC/C are shown below the panels.



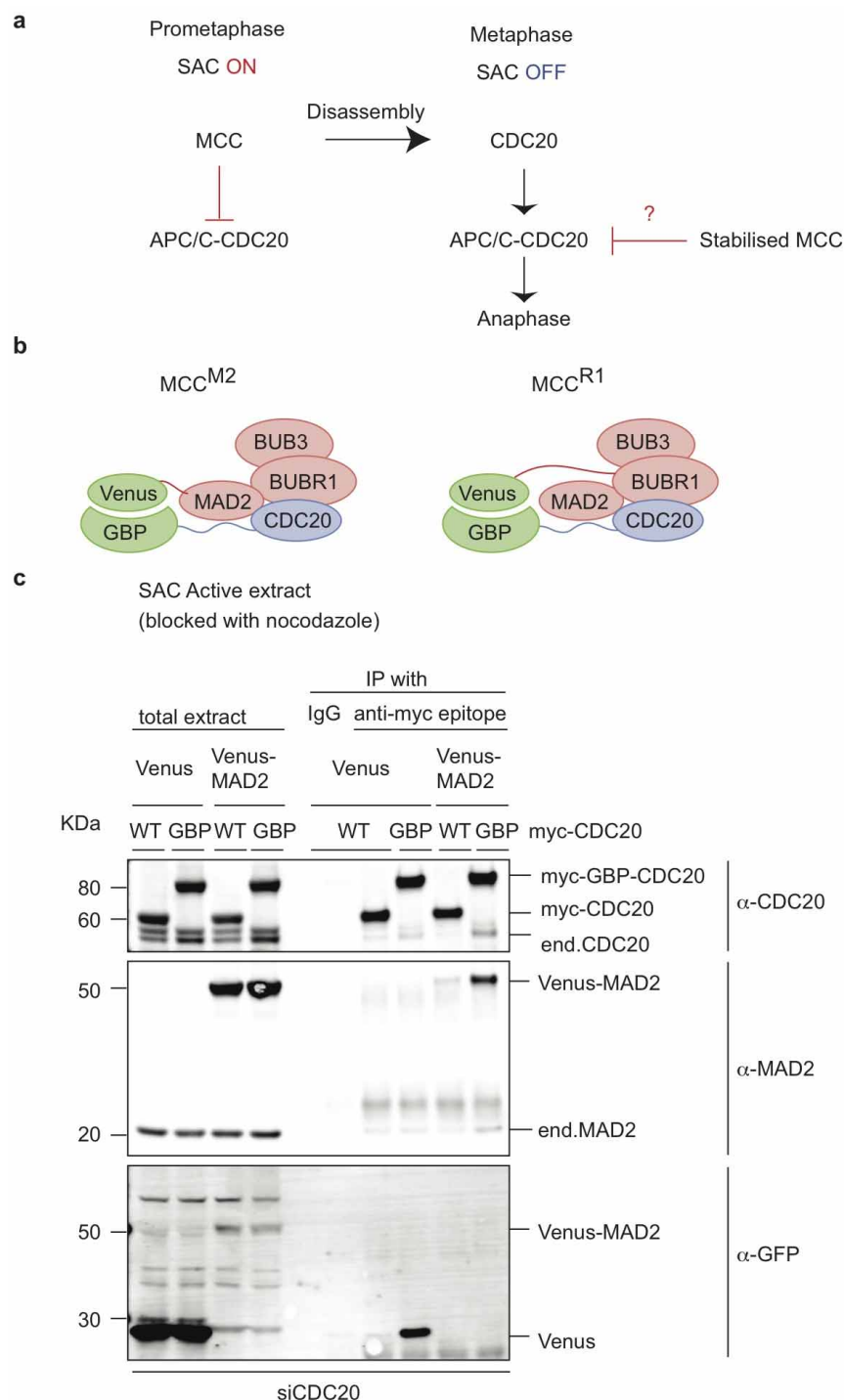
Extended Data Figure 4 | Cells expressing the D-box and KEN box receptor mutants of CDC20 can degrade cyclin B1 in nocodazole. Cyclin B1-Venus degradation was analysed in siRNA CDC20-treated cells rescued with siRNA-resistant versions of 3×Flag-CDC20, wild-type, or ΔDR, or ΔKR

mutants, in the presence of nocodazole (0.33 μM). The fluorescence of individual cells was measured, the value at NEBD set to 1 and the mean ± s.e.m. for all cells plotted. *n*, number of cells analysed from at least two independent experiments.



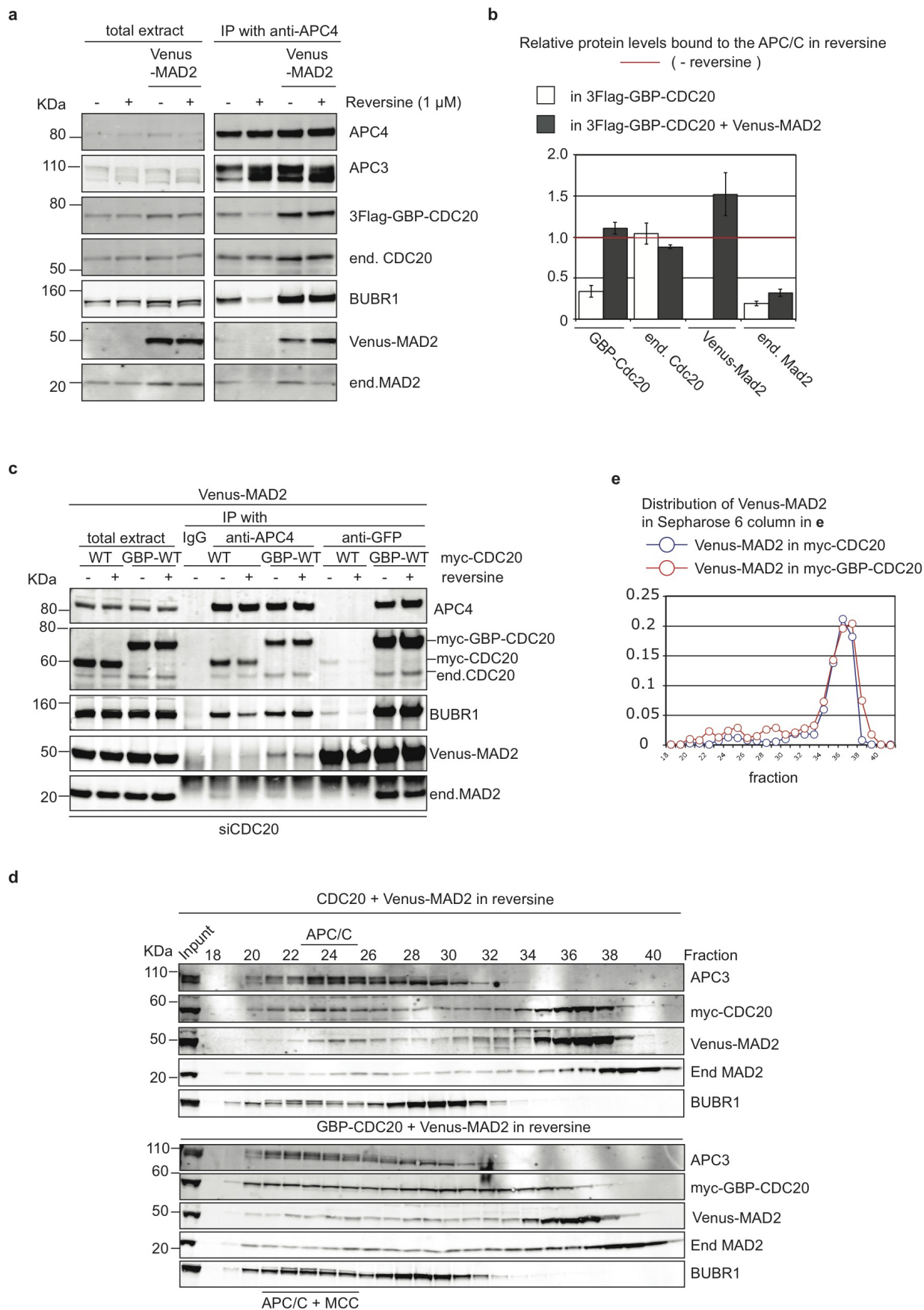
Extended Data Figure 5 | Characterization of the MCC containing D-box or KEN box 2 mutants of BUBR1. **a**, Core rMCC assembled with ^{SBP}BUBR1 wild type, or ΔD-box, or ΔKEN2 mutants, was purified as in Extended Data Fig. 1a, b and analysed on a LiCOR Odyssey scanner at 680 nm after SDS-PAGE and Coomassie blue R250 staining. **b**, The core rMCC mutants prepared in **a** were assayed as APC/C inhibitors in an *in vitro* ubiquitylation assay as in Fig. 1d. **c**, Insect cell extracts expressing CDC20 with ^{SBP}BUBR1, either wild type, or ΔD-box or ΔKEN2 mutants, were incubated with streptavidin beads. The proteins retained on the streptavidin beads were analysed by quantitative immunoblotting. Results in panels **a–c** are representative of two independent biological replicates. **d**, HeLa cells were

treated with siRNA against BUBR1 and rescued with 3×Flag–Cerulean–BUBR1, either wild-type or the ΔD-box mutant, and mitosis analysed in 0.116 μM Taxol as in Fig. 3d. The time from NEBD to anaphase (or mitotic exit) was measured and plotted as a box and whisker chart. *n*, number of analysed cells from two independent biological replicates. **e**, HeLa cells were treated with siRNA against BUBR1 and rescued with siRNA resistant 3×Flag–Cerulean–BUBR1, either wild type or the ΔD-box mutant, then analysed by immunostaining. Cells were stained with anti-Flag M2 and anti-ACA antibodies, and Hoechst 33342, and representative images of prometaphase cells from two independent biological replicates are shown. Scale bar, 10 μm.



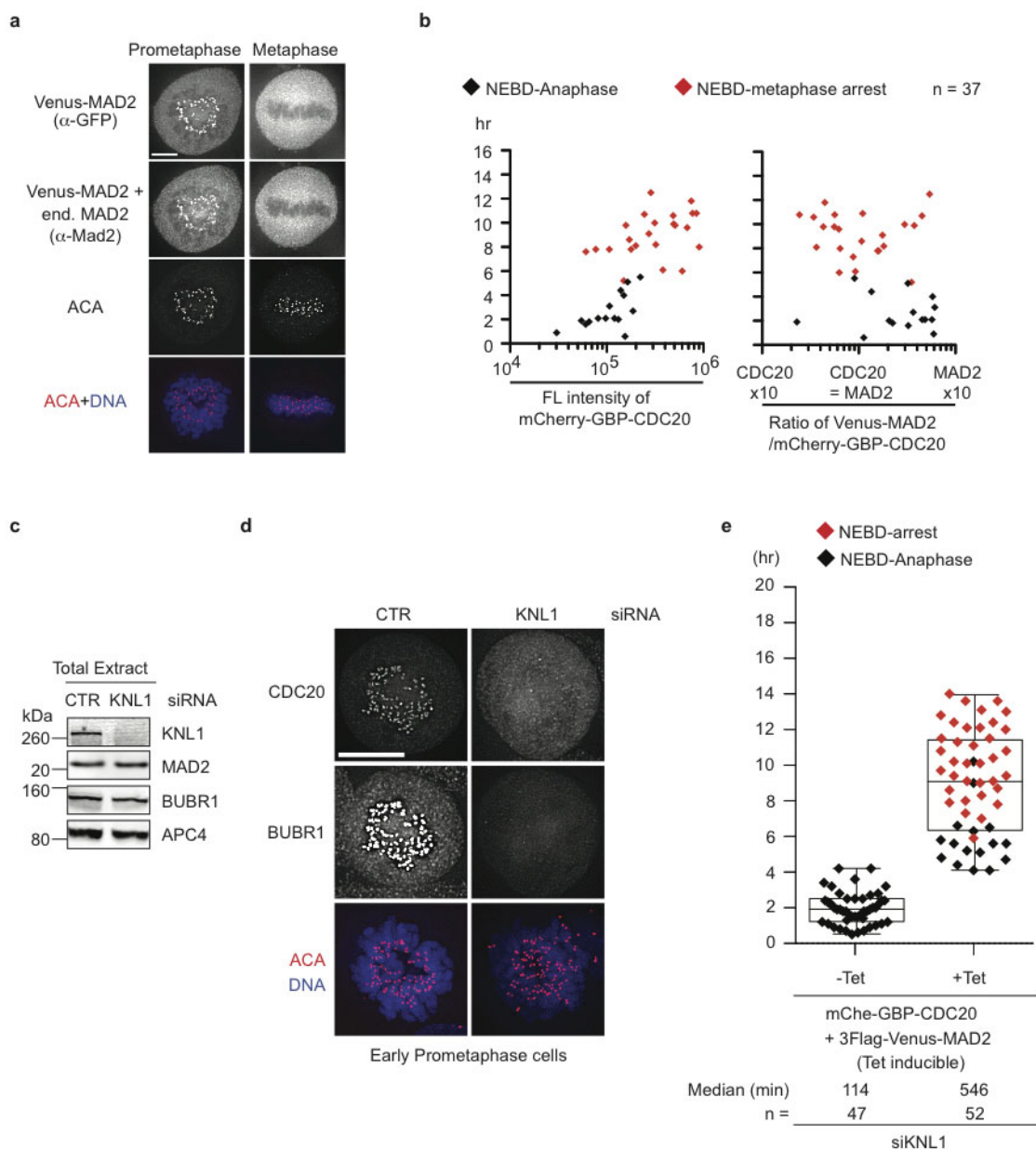
Extended Data Figure 6 | Stabilizing the interaction between MAD2 and CDC20. **a**, Schematic of how a stabilized MCC might block cells in metaphase. At prometaphase, when the SAC is 'ON', CDC20 is inhibited both by incorporation into the MCC and through binding to the MCC. At metaphase when the SAC is 'OFF', CDC20 is released from the MCC and activates the APC/C, which results in an anaphase delay. **b**, Schematic of a stabilized MCC. To stabilize the MCC we took advantage of the binding between yellow fluorescent protein (Venus) and GFP-binding domain (GBP), which is a 13kDa domain from a camelid antibody that binds strongly and specifically to GFP and YFP¹⁸. MAD2 and BUBR1 were tagged with Venus and the GBP domain

was tagged to CDC20. We refer to the MCC containing a stabilized MAD2-CDC20 interaction as MCC^{M2}, and that with stabilized BUBR1-CDC20 as MCC^{R1}. **c**, GBP- and Venus-fusion proteins bind stably to each other *in vivo*. HeLa cell lines expressing siRNA-resistant myc-CDC20 or myc-GBP-CDC20 were transfected with plasmids encoding either Venus alone or Venus-MAD2, followed by siRNA treatment against CDC20. After a single thymidine block and release, the cells were arrested at prometaphase by treating with nocodazole, and harvested by mitotic shake-off 48 h after the siRNA treatment. Proteins were immunoprecipitated with anti-myc epitope antibodies before analysis by quantitative immunoblotting with the indicated antibodies. WT, myc-CDC20; GBP, myc-GBP-CDC20. Results in panel **c** are representative of three independent biological replicates.



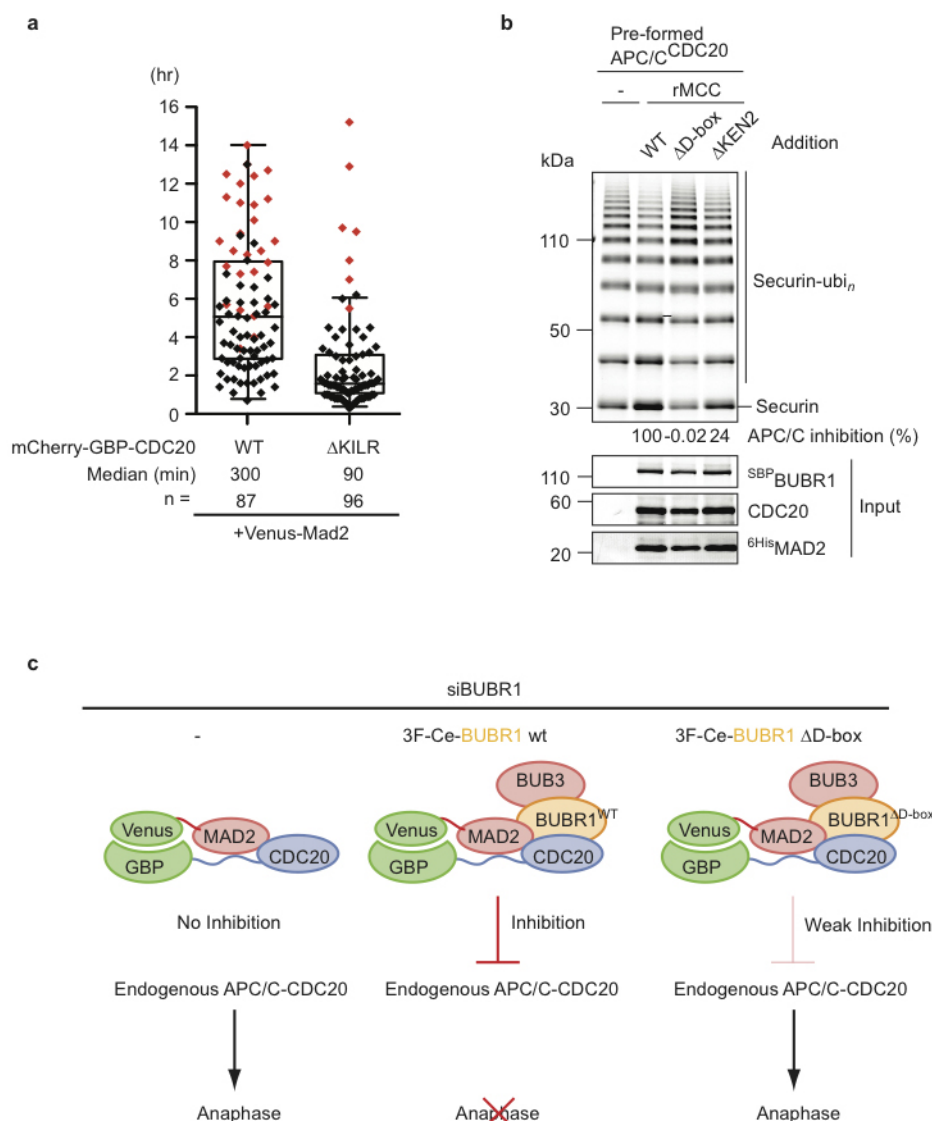
Extended Data Figure 7 | Stabilizing the interaction between MAD2 and CDC20 prevents disassembly of the MCC *in vivo*. **a, b**, Tethering CDC20 to MAD2 prevents MCC disassembly and release from the APC/C. **a**, Empty plasmids or plasmids encoding Venus–MAD2 were transfected into HeLa cell lines expressing 3×Flag–GBP–CDC20 and the cells synchronized at prometaphase by thymidine release followed by a nocodazole block. Cells were harvested by mitotic shake off and separated into two cultures after washing once in medium. One culture was harvested immediately (–reversine) and the other resuspended in medium containing 1 μ M reversine and 10 μ M MG132 (+reversine) for 1 h before harvesting. The APC/C was immunoprecipitated with an anti-APC4 antibody and the immunoprecipitates analysed by quantitative immunoblotting. We note that the APC/C preferred to bind endogenous CDC20 over GBP–CDC20 as the co-activator *in vivo* (see +reversine lane in control cells) but the MCC^{M2} did not sequester endogenous CDC20 from the APC/C (see +reversine lane in GBP–CDC20 + Venus–MAD2 cells). **b**, Mean \pm s.e.m. of the relative amounts of the indicated proteins in the APC4 immunoprecipitates calculated from four independent biological experiments. The amount of protein bound to the APC/C in the absence of reversine was set to 1 (red line). **c–e**, Tethering CDC20 to MAD2 prevents MCC

disassembly and release from the APC/C in the absence of endogenous CDC20. **c**, Plasmids encoding Venus–MAD2 were transfected into HeLa cell lines expressing the indicated CDC20 fusion proteins following siRNA treatment against CDC20 for 48 h. Cells were synchronized at prometaphase then treated with reversine, and anti-APC4 and anti-GFP immunoprecipitates were analysed as in **a**. WT, myc–CDC20; GBP, myc–GBP–CDC20. Note that endogenous CDC20 could not be inhibited through exchange into MCC^{M2} because a core MCC composed of Venus–MAD2 and untagged CDC20 disassembled. **d**, HeLa cell lines expressing myc–CDC20 (upper blots) or myc–GBP–CDC20 (lower blots) were transfected with a plasmid encoding Venus–MAD2 followed by siRNA treatment against CDC20 for 48 h. Cells were synchronized at prometaphase and treated with reversine as indicated in **a**. Total cell extracts were analysed by size exclusion chromatography on a Sepharose 6 column and fractions were analysed by quantitative immunoblotting against the indicated proteins and the relative amounts of Venus–MAD2 plotted in panel **e** with the sum of Venus–MAD2 intensities set to 1. The migration of APC/C or APC/C–MCC is annotated below panel **d**. All results are representative of three independent biological replicates.



Extended Data Figure 8 | KNL1 (also known as CASC5) is not required for a stabilized MCC to inhibit anaphase. **a**, HeLa cells expressing MCC^{M2} in Fig. 4a were analysed by immunostaining. The cells were stained with anti-GFP, anti-MAD2, anti-ACA and Hoechst 33342, and representative images of prometaphase and metaphase cells from two independent biological replicates are shown. Scale bar, 5 μ m. **b**, The time from NEBD to anaphase in Fig. 4a was plotted against the intensity of mCherry-GBP-CDC20 (left) or the ratio of Venus-MAD2 to mCherry-GBP-CDC20 (right). The ratio of Venus-MAD2 to mCherry-GBP-CDC20 was calibrated by measuring fluorescence intensity of a mCherry-GBP-Venus fusion protein in HeLa cells. **c–e**, MCC^{M2} delays anaphase when KNL1 is depleted. **c**, HeLa cells were treated with siRNA against

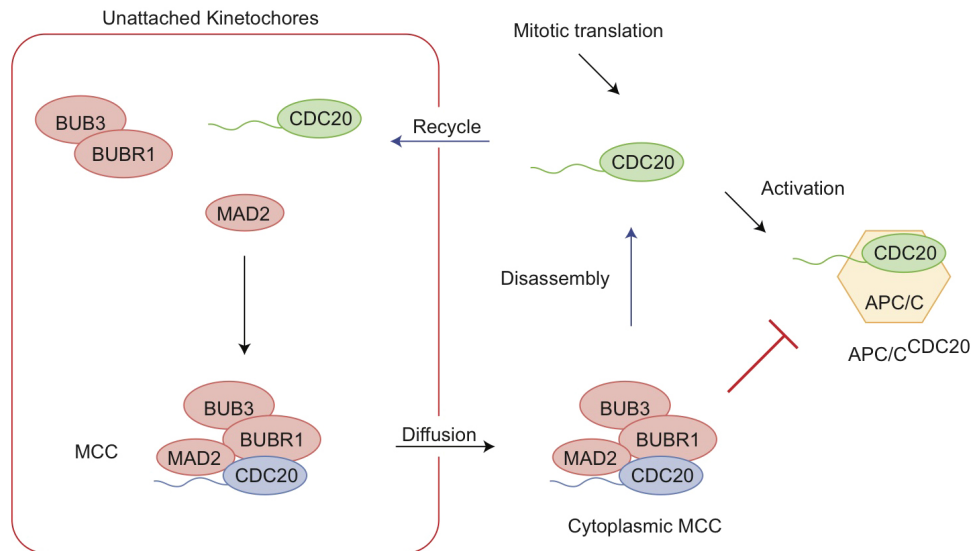
KNL1 for 72 h and total cell extracts were analysed by quantitative immunoblotting with the indicated antibodies. **d**, HeLa cells treated as in **c** were analysed by immunostaining. The cells were stained with anti-CDC20, anti-BUBR1, anti-ACA and Hoechst 33342, and representative images of early prometaphase from two independent biological replicates are shown. Scale bar, 10 μ m. **e**, HeLa cell lines stably expressing mCherry-GBP-CDC20 and an inducible 3 \times Flag-Venus-MAD2 (expressed from a tetracyclin-inducible promoter) were treated with siRNA against KNL1 as in **c**. Progression through mitosis was analysed in the presence (+Tet) or absence (–Tet) of tetracyclin, and analysed as in Fig. 4b. *n*, number of cells from two independent biological replicates.



Extended Data Figure 9 | Functional MCC^{M2} is required to delay anaphase.

a, HeLa cells were transfected with plasmids encoding Venus-MAD2 and either wild-type or a MAD2-binding defective (Δ KILR) mutant of CDC20 tagged with mCherry-GBP, and mitotic progression was analysed as in Fig. 4a. *n*, number of cells from three independent biological replicates. **b**, The core rMCC mutants used in Extended Data Fig. 5a were incubated with preformed APC/C^{CDC20} and assayed as APC/C inhibitors in an *in vitro* ubiquitylation assay as in Fig. 1e. The extent of APC/C inhibition (incubation of MCC^{WT} set

to 1.0) is shown below the securin panel. This result is representative of two independent experiments. **c**, Schematic of the inhibitory activities of the stabilized MCCs in BUBR1-depleted cells used in Fig. 4c. When BUBR1 is depleted, MAD2 and CDC20 cannot form the MCC to inhibit endogenous CDC20 (left). When rescued with wild-type BUBR1, MCC^{M2} can form and inhibit endogenous CDC20 to delay anaphase. By contrast, when rescued by the BUBR1 Δ D-box mutant, MCC^{M2} can only weakly inhibit endogenous CDC20 and cells can proceed into anaphase.



Extended Data Figure 10 | Model for how the MCC could disseminate the 'wait anaphase' signal. Unattached kinetochores catalyse MCC formation and the MCC disseminates the 'wait anaphase' signal through the cytoplasm (black arrows). When the MCC disassembles (blue arrows), this releases

CDC20, which along with newly synthesized CDC20, can have two fates: to be recruited to unattached kinetochores and incorporated into the MCC, or to bind the APC/C to form APC/C^{CDC20}. The MCC is able to inhibit both unbound CDC20 and CDC20 bound to the APC/C (red bars).

Uncovering the polymerase-induced cytotoxicity of an oxidized nucleotide

Bret D. Freudenthal¹, William A. Beard¹, Lalith Perera¹, David D. Shock¹, Taejin Kim^{2,3}, Tamar Schlick^{2,3} & Samuel H. Wilson¹

Oxidative stress promotes genomic instability and human diseases¹. A common oxidized nucleoside is 8-oxo-7,8-dihydro-2'-deoxyguanosine, which is found both in DNA (8-oxo-G) and as a free nucleotide (8-oxo-dGTP)^{2,3}. Nucleotide pools are especially vulnerable to oxidative damage⁴. Therefore cells encode an enzyme (MutT/MTH1) that removes free oxidized nucleotides. This cleansing function is required for cancer cell survival^{5,6} and to modulate *Escherichia coli* antibiotic sensitivity in a DNA polymerase (pol)-dependent manner⁷. How polymerases discriminate between damaged and non-damaged nucleotides is not well understood. This analysis is essential given the role of oxidized nucleotides in mutagenesis, cancer therapeutics, and bacterial antibiotics⁸. Even with cellular sanitizing activities, nucleotide pools contain enough 8-oxo-dGTP to promote mutagenesis^{9,10}. This arises from the dual coding potential where 8-oxo-dGTP(*anti*) base pairs with cytosine and 8-oxo-dGTP(*syn*) uses its Hoogsteen edge to base pair with adenine¹¹. Here we use time-lapse crystallography to follow 8-oxo-dGTP insertion opposite adenine or cytosine with human pol β , to reveal that

insertion is accommodated in either the *syn*- or *anti*-conformation, respectively. For 8-oxo-dGTP(*anti*) insertion, a novel divalent metal relieves repulsive interactions between the adducted guanine base and the triphosphate of the oxidized nucleotide. With either templating base, hydrogen-bonding interactions between the bases are lost as the enzyme reopens after catalysis, leading to a cytotoxic nicked DNA repair intermediate. Combining structural snapshots with kinetic and computational analysis reveals how 8-oxo-dGTP uses charge modulation during insertion that can lead to a blocked DNA repair intermediate.

A primary defence mechanism against oxidative DNA damage is base excision repair, which in eukaryotes utilizes pol β ^{12,13}. During times of oxidative stress, pol β can perform futile repair by inserting 8-oxo-dGTP opposite cytosine (Cy) or adenine (Ad) (Fig. 1a–c), and it is implicated in driving tumorigenesis^{14,15}. Pol β binds to gapped DNA in an open conformation; upon binding the incoming nucleotide it undergoes a conformational change to form the pre-catalytic closed complex with two active site divalent metals ions: the catalytic (Mg_c) and nucleotide

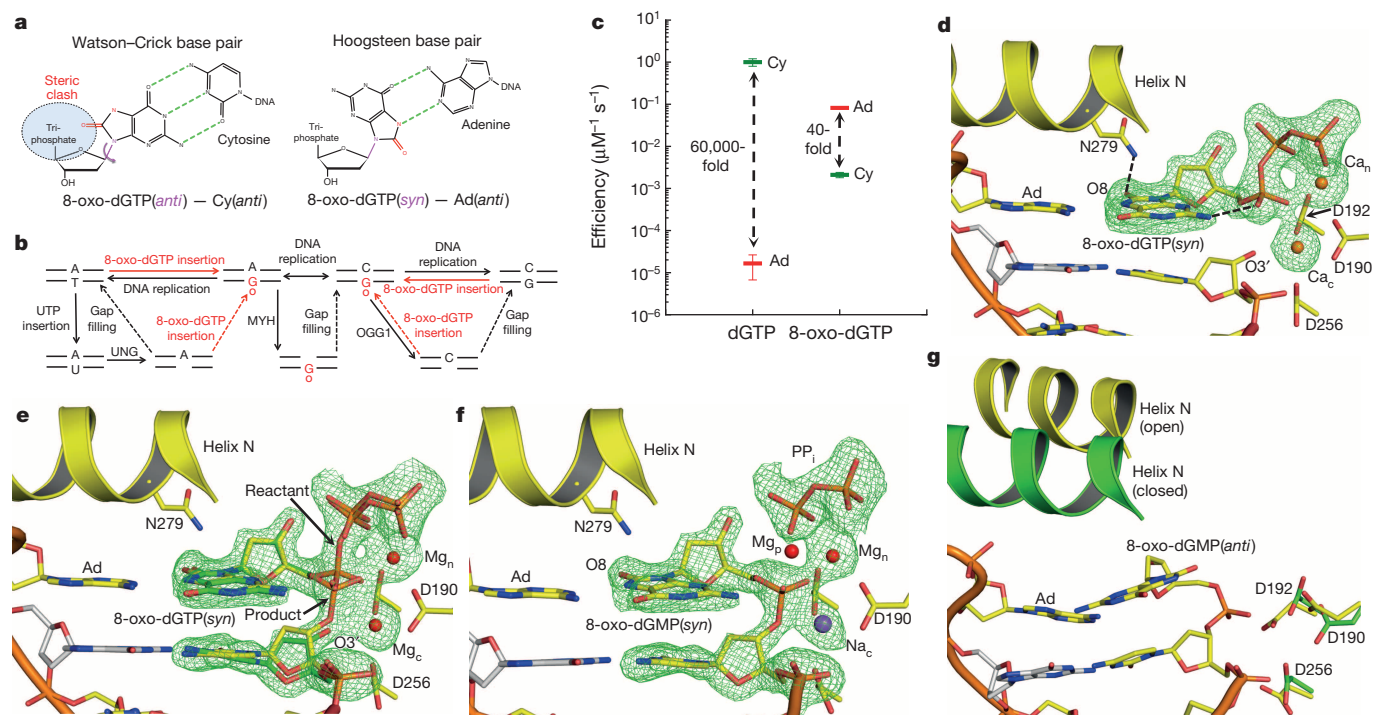


Figure 1 | 8-Oxo-dGTP specificity and insertion opposite adenine.

a, 8-Oxo-dGTP base pairing with Cy or Ad. **b**, Pathways associated with 8-oxo-G DNA repair for the A to C transversion. Dashed lines are pol β insertion events. **c**, Pol β insertion efficiency of 8-oxo-dGTP and dGTP opposite either templating Cy or Ad. Discrimination for the preferred nucleotide is indicated. **d**, 8-Oxo-dGTP(*syn*):Ad pre-catalytic complex. **e**, Phosphodiester bond formation after a 20 s soak in $MgCl_2$ with the reactant (green) and product

(yellow) states. **f**, Closed 8-oxo-dGMP:Ad product complex after a 40 s soak in $MgCl_2$. **g**, Open 8-oxo-dGMP:Ad product complex after a 90 s soak. Closed conformation is shown in green (Protein Data Bank accession number 2FMS). $F_o - F_c$ maps (3σ) are in green. Ca^{2+} , Mg^{2+} , and Na^+ are orange, red, and purple spheres respectively. The catalytic, nucleotide, and product metals are denoted with subscripts c, n, and p, respectively.

¹Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, PO Box 12233, Research Triangle Park, North Carolina 27709-2233, USA. ²Department of Chemistry, New York University, and NYU-ECNU Center for Computational Chemistry at NYU Shanghai, 10th Floor Silver Center, 100 Washington Square East, New York, New York 10003, USA. ³Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012, USA.

(Mg_n) metals^{16–18}. This complex is optimized for nucleotidyl transfer, forming pyrophosphate (PP_i), and following catalysis, pol β reopens, releasing PP_i.

Soaking open binary crystals of pol β bound to DNA containing a templating Ad in a cryosolution with 8-oxo-dGTP and CaCl₂ results in a closed pre-catalytic ground state ternary complex (Extended Data Table 1). The incoming 8-oxo-dGTP(*syn*) Hoogsteen base pairs with Ad (Fig. 1d). The active site is in a similar conformation as previously observed using a dideoxy-terminated primer (Extended Data Fig. 1a)¹⁹, with a change in the primer terminus sugar pucker to C3'-endo (Extended Data Table 2). Unique active site interactions include Asn 279 hydrogen bonding to O8 of 8-oxo-dGTP(*syn*) and an intramolecular hydrogen bond between N2 and the pro-S_p oxygen on Pα of 8-oxo-dGTP(*syn*) (Fig. 1d). This is consistent with previous studies identifying a role of Asn 279 in stabilizing 8-oxo-dGTP(*syn*)²⁰.

To observe catalysis the ground state crystals were transferred to a solution containing MgCl₂ for 20 s (Extended Data Table 1). Density corresponding to both the reactant and product was observed (Fig. 1e and Extended Data Fig. 1b), and on the basis of occupancy refinement the reaction was 60% 'complete'. Compared with the ground state there is only moderate movement in the active site, at Pα and O3' (Extended Data Fig. 1b, c). The Hoogsteen base pairing and hydrogen bonding interactions that stabilize the planar *syn*-conformation are maintained, and the enzyme remains in the closed conformation. These observations are consistent with 8-oxo-dGTP insertion opposite Ad exhibiting a high catalytic efficiency (Fig. 1c).

Extending the MgCl₂ soak time to 40 s resulted in complete turnover of the reactant 8-oxo-dGTP(*syn*) in the closed polymerase conformation (Extended Data Table 1 and Fig. 1f). This post-chemistry complex shows the Hoogsteen base pairing and the intramolecular hydrogen bond at N2 are maintained, and Asn 279 hydrogen bonds with O8. The catalytic Mg²⁺ has been replaced by Na⁺, while the nucleotide Mg²⁺ remains in the active site coordinating PP_i (Extended Data Fig. 1d, e). The closed product complex contains a new Mg²⁺ product metal (Mg_p) that bridges the backbone phosphate of 8-oxo-dGMP(*syn*) and PP_i.

Extending the MgCl₂ soak to 90 s results in a closed-to-open conformational change (Fig. 1g). The PP_i and associated metals have disassociated, and the inserted 8-oxo-dGMP(*anti*) has lost the interaction between Asn 279 and O8, promoting destabilization of the Hoogsteen base pairing. The inserted 8-oxo-dGMP has a high B-factor (57 Å²; Extended Data Table 1), with the most stable position displaced into the major groove and a weak hydrogen bond formed between N6 of Ad and N3 of 8-oxo-dGMP (Extended Data Fig. 1f).

The insertion efficiency of 8-oxo-dGTP opposite Cy is much less than opposite Ad (Fig. 1c); this may arise from a clash between O8 and Pα of 8-oxo-dGTP(*anti*) (Fig. 1a)^{19,21}. To probe how this clash is accommodated, we soaked open binary pol β complex crystals with a templating Cy in a cryosolution containing 8-oxo-dGTP/CaCl₂. The resulting closed ternary ground state complex contains 8-oxo-dGTP(*anti*) Watson–Crick base pairing with the templating Cy (Extended Data Table 3 and Fig. 2a). The clash at O8 is partly eased by an altered sugar pucker, glycosidic angle, buckle, and shear compared with dGTP(*anti*) (Extended Data Table 2).

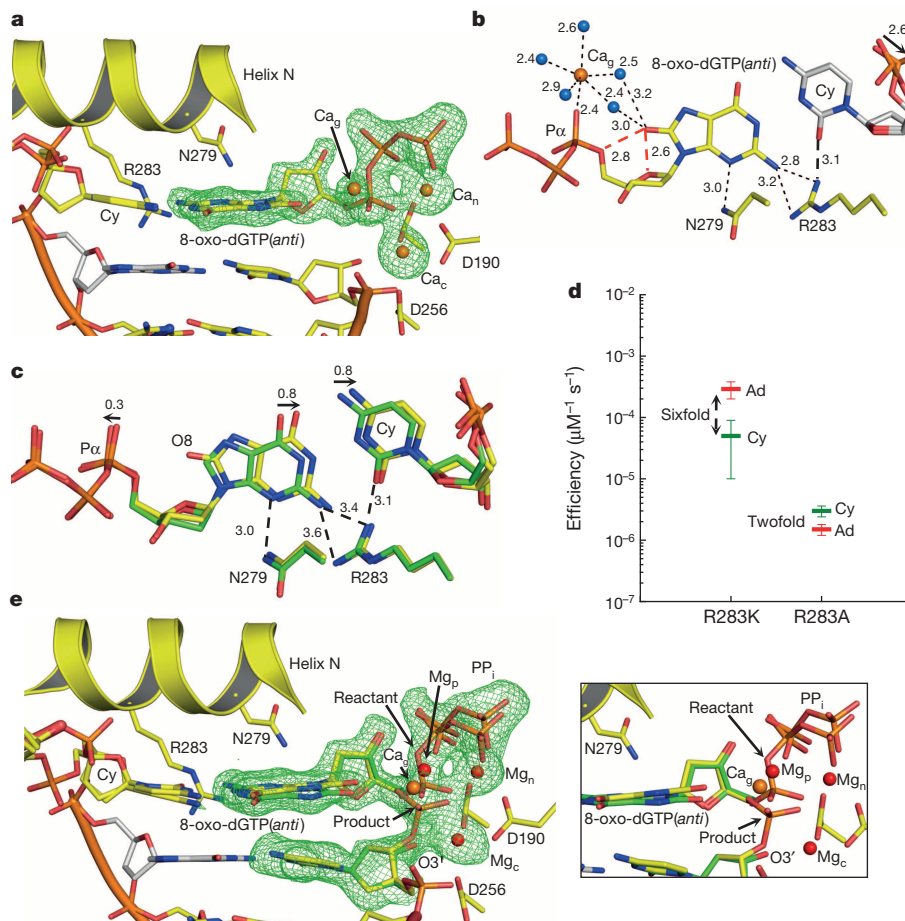


Figure 2 | 8-Oxo-dGTP insertion opposite cytosine. **a**, 8-Oxo-dGTP: Cy pre-catalytic active site. **b**, A 90° rotation relative to **a**. Coordinating waters (blue) with distances (in ångströms) are indicated. Clash at O8 is shown with red dashes. The Cy backbone shift is indicated. **c**, Overlay of 8-oxo-dGTP(*anti*) and dGTP(*anti*) opposite Cy is shown in yellow or green, respectively. The shift from the adducted O8 is indicated. **d**, Catalytic efficiency of 8-oxo-dGTP

insertion opposite Cy or Ad for pol β R283K and R283A. **e**, The active site after a 40 s soak in MgCl₂ with the reactant (green) and product (yellow) species shown. A close-up without density is in the adjacent panel. $F_o - F_c$ maps (3σ) are in green. Ca²⁺, Mg²⁺, and Na⁺ are orange, red, and purple spheres, respectively. The catalytic, nucleotide, ground, and product metals are denoted with a subscript c, n, g, and p, respectively.

The incoming 8-oxo-dGTP(*anti*) N3 and N2 atoms are within hydrogen-bonding distance of Asn 279 and Arg 283, respectively (Fig. 2b). To determine whether these are unique contacts, we solved the structure of dGTP(*anti*) Watson–Crick base pairing with Cy in the presence of CaCl_2 (Extended Data Table 3). Comparing these two structures indicates the contact with Arg 283 may be unique to the incoming 8-oxo-dGTP(*anti*) because O8 causes the triphosphate and base to move 1.1 Å apart (Fig. 2c). This shift promotes the phosphate backbone of Cy to move 2.6 Å into the minor groove (Fig. 2b). Mutating Arg 283 to lysine or alanine reduced the 8-oxo-dGTP specificity from favouring 8-oxo-dGTP(*syn*) insertion opposite Ad by 40-fold for wild-type, to sixfold and twofold for the R283K and R283A mutants, respectively (Fig. 2d). This loss of discrimination against insertion of 8-oxo-dGTP(*anti*) opposite Cy indicates Arg 283 promotes the mutagenic insertion of 8-oxo-dGTP(*syn*) by acting as a steric gate to prevent insertion opposite Cy. Interestingly, during the bypass of 8-oxo-G in the templating position, Arg 283 stabilizes the mutagenic *syn*-conformation²².

The changes described above do not fully alleviate the clash between O8 and the sugar–phosphate backbone of the incoming 8-oxo-dGTP(*anti*) (red dashes in Fig. 2b). Surprisingly, this clash is accommodated by an additional divalent metal (Ca_g) observed coordinating the pro- S_p oxygen of P_α and five water molecules (Fig. 2a, b). Two of the water molecules are also within hydrogen bonding distance of O8 (Fig. 2b). To verify the presence of this additional divalent metal-binding site near P_α , we soaked the closed pre-catalytic ground state complex (8-oxo-dGTP: Cy) in MnCl_2 for 5 s (Extended Data Table 3). This allowed metal exchange before catalysis had appreciatively occurred (Extended Data Fig. 2a). The presence of Mn^{2+} in the catalytic, nucleotide, and ground state sites was

verified by anomalous density. Overlaying the ground state complex with Ca^{2+} and Mn^{2+} verified the ground state metal site binds divalent cations (Extended Data Fig. 2b).

To observe 8-oxo-dGTP(*anti*) insertion we soaked a ground state crystal in MgCl_2 for 40 s (Extended Data Table 3). Pol β remained in the closed conformation with density corresponding to both product and reactant species (Fig. 2e). Occupancy refinement indicates the reaction is 40% complete. The Watson–Crick base pairing interactions are maintained with only moderate movement at the reacting atoms (P_α and $\text{O}3'$; Fig. 2e). The phosphate backbone of the templating Cy is observed in two conformations corresponding to reactant and product. Surprisingly, density for both ground state and product metals can be observed, indicating there are two distinct populations within the crystal (Extended Data Fig. 3).

Soaking pre-catalytic complex crystals in MgCl_2 for 60 s results in a closed product complex (Extended Data Table 3). 8-Oxo-dGTP(*anti*) has been completely inserted with the sugar pucker shifting from a $\text{C}4'$ -exo (ground state) to $\text{C}3'$ -endo (Extended Data Table 2 and Fig. 3a). Importantly, there is no density corresponding to the ground state metal and only the product-associated metal is present (Fig. 3a, b). The clash at O8 has not been fully alleviated in the product complex and is probably being mediated by the product Mg^{2+} (Fig. 3c). Overlaying the closed ground (0 s) and product states (60 s) indicates that ground state and product metal-binding sites are in distinct positions separated by 2.0 Å that are dependent on having either substrate or product present (Fig. 3d). Extending the soak time in MgCl_2 to 120 s resulted in pol β transitioning to an open conformation (Extended Data Table 3). Watson–Crick base pairing is lost, 8-oxo-dGMP stacks over the templating base with

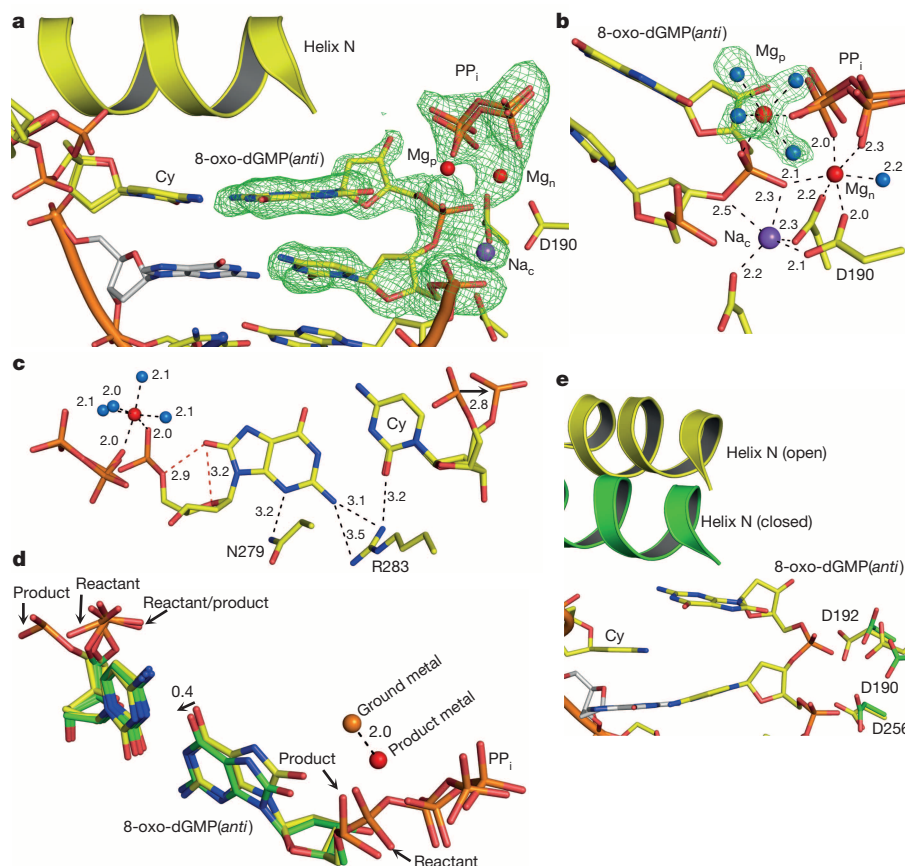


Figure 3 | Product complex with 8-oxo-dGMP(*anti*):cytosine. **a**, Closed 8-oxo-dGMP: Cy product complex after a 60 s soak in MgCl_2 . **b**, Active site of the closed 8-oxo-dGMP: Ad product complex with coordination distances (in ångströms) and waters (blue). **c**, A 90° rotation relative to **a**. The clash at O8 is shown with red dashed lines. **d**, Overlay of the ground (green) and product (yellow) 8-oxo-dGTP(*anti*): Cy complexes. The distance (in ångströms)

between the ground and product metal sites is shown with a dashed line. **e**, Open 8-oxo-dGMP: Cy product complex after a 120 s soak in MgCl_2 . Closed conformation is shown in green (Protein Data Bank accession number 2FMS). $F_o - F_c$ maps (3σ) are in green. Ca^{2+} , Mg^{2+} , and Na^+ are orange, red, and purple spheres, respectively.

a high B-factor (65.4 \AA^2), and the remaining PP_i and associated metals have dissociated (Fig. 3e).

Quantum mechanical analysis allows the point charge on each atom of dGTP(*anti*) and 8-oxo-dGTP(*anti*) to be calculated (Extended Data Table 4 and Extended Data Fig. 4). Figure 4a illustrates the charge difference between 8-oxo-dGTP and dGTP mapped onto each atom of dGTP. The adducted O8 causes the oxygen bridging the sugar moiety and triphosphate (O5) to become more positive. Likewise, the pro- S_p oxygen of $\text{P}\alpha$ becomes more negative and may facilitate recruitment of the ground state metal. We also calculated the charge on each atom of 8-oxo-dGTP(*anti*) with all three metals (Extended Data Table 4 and Extended Data Fig. 4). Figure 4b shows the difference in charge of 8-oxo-dGTP(*anti*) with three metals from that with two metals (Ca_c and Ca_n) mapped onto the structure of 8-oxo-dGTP(*anti*). The largest electronegative change is localized on the triphosphate at key catalytic atoms. This includes making $\text{P}\alpha$ and $\text{P}\beta$ more positive, while their bridging oxygen becomes more negative.

Molecular dynamics simulations using the pre-catalytic 8-oxo-dGTP(*anti*) or dGTP(*anti*) opposite Cy structures indicate a stable Mg_g coordination sphere (Extended Data Table 5). With 8-oxo-dGTP(*anti*), one of the Mg_g -coordinating water molecules forms a hydrogen bond with O8, while the inter-atomic distances in the active site maintain proper catalytic values (Extended Data Fig. 5a, b). In comparison, the dGTP(*anti*) system exhibits poor geometry. The average distance of $\text{P}\alpha$ - $\text{O}3'$ increases to $5.25 \pm 1.0 \text{ \AA}$ (Extended Data Fig. 5c), and the coordination between Mg_c and $\text{O}3'$ is broken owing to a newly established coordination network with another water molecule at 40 ns. All these events in the dGTP(*anti*) system induce larger root mean squared deviations than those of the 8-oxo-dGTP(*anti*) system when the evolving molecular dynamics system is compared with the initial structure, indicating that Mg_g in the dGTP(*anti*) is less favourable and associated with a much less competent-for-chemistry geometry compared with the 8-oxo-dGTP(*anti*) system (Extended Data Fig. 5c, d).

Interestingly, we captured 8-oxo-dGTP(*anti*) Watson–Crick base pairing in a nearly identical manner to that observed for dGTP(*anti*). This required recruitment of an additional divalent metal near $\text{P}\alpha$ of 8-oxo-dGTP(*anti*) that forms a stable hydration shell within hydrogen bonding

distance to the adducted oxygen. Consequently, this metal helps alleviate the clash at O8 and permits binding of 8-oxo-dGTP(*anti*) without repositioning $\text{P}\alpha$. A similar phenomena is observed in the 8-oxo-dGTP(*syn*) conformation, where the exocyclic N2 forms an intramolecular hydrogen bond with $\text{P}\alpha$, stabilizing good nascent base pair geometry. This implies that the interaction of $\text{P}\alpha$ by either N2 or a divalent metal cation is a hallmark characteristic of 8-oxo-dGTP insertion during both mutagenic and non-mutagenic insertion, respectively.

Damaged substrates complicate the ability of DNA polymerases to select the correct nucleotide^{23–25}. Structural studies have identified that correct and incorrect non-damaged nucleotides are discriminated from each other on the basis of proper alignment of catalytic atoms. The planar nature of 8-oxo-dGTP places $\text{O}3'$ of the primer terminus near $\text{P}\alpha$ so that only minor structural rearrangements are needed for nucleotidyl transfer. Insertion of a correct nucleotide results in a stable ternary product complex, while incorrect insertion promotes rapid re-opening of the enzyme²⁶. Immediately following 8-oxo-dGTP insertion the polymerase reopened, similar to an incorrect insertion and implying 8-oxo-dGMP promotes instability. These findings show that 8-oxo-dGTP utilizes characteristics of both correct and incorrect insertion elements.

Recent structural studies have identified a transient third metal-binding site associated with the products^{26,27}. For pol β , this third metal (Mg_p) was only observed in the product complex following insertion of the correct, but not incorrect, nucleotide. The appearance of the product metal was unexpected following 8-oxo-dGTP(*syn*) misinsertion opposite Ad, but consistent with the good base pair geometry exhibited by this mispair. It appears that these adjunct metal sites are necessary to neutralize negative charge that may be inherent in the substrate or that transiently develops during chemistry.

We can infer a mechanistic model for the role of these adjunct metal ions during catalysis (Fig. 4c). In the pre-catalytic ground state the primer terminus, Mg_c , Mg_n , and incoming nucleotide are bound. De-protonation of $\text{O}3'$ initiates nucleophilic attack at $\text{P}\alpha$ and as $\text{O}3'$ approaches $\text{P}\alpha$ it sterically clashes with the non-bridging oxygens of $\text{P}\alpha$. This results in a transition state and localized charges on $\text{O}3'$, $\text{P}\alpha$, and $\text{O}_{\text{P}\alpha}$ that recruit a metal ion to polarize $\text{P}\alpha$, thus facilitating $\text{O}3'$ attack by making $\text{P}\alpha$ more positive and $\text{O}_{\text{P}\alpha-\beta}$ more negative. Such a role has been postulated for a

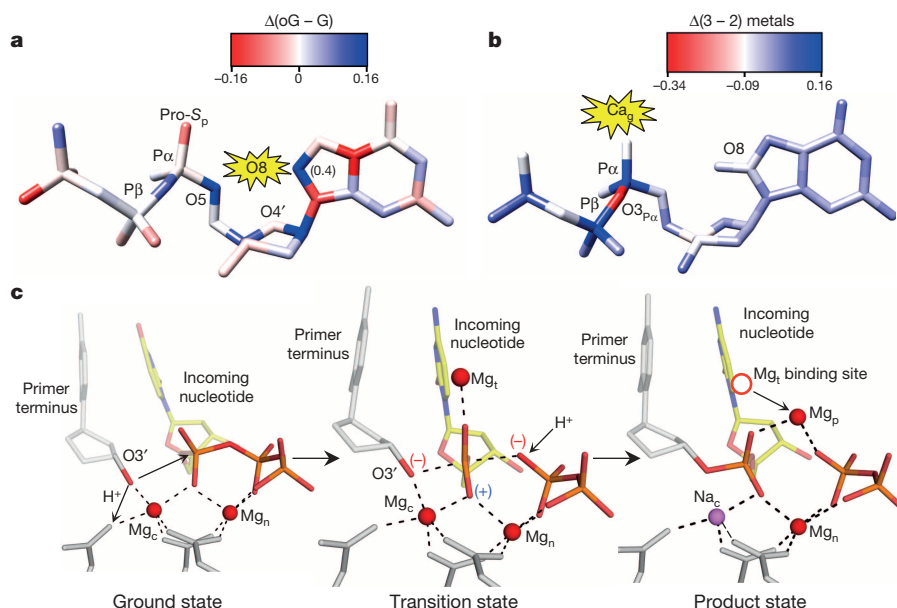


Figure 4 | Charge modulation of the polymerase active site. **a**, The charge difference for each atom of dGTP(*anti*) and 8-oxo-dGTP(*anti*) is plotted onto dGTP with a colour key shown. The only value that does not fall within the indicated range is C8 ($0.4e^-$). **b**, The charge difference for each atom of 8-oxo-dGTP(*anti*) with three and two metals is mapped onto 8-oxo-dGTP with a colour key shown. **c**, Proposed model for the catalytic mechanism of

nucleotide insertion. The primer terminus is shown in grey with $\text{O}3'$ in red and the incoming nucleotide in yellow. The transition state includes the catalytic (Mg_c), nucleotide (Mg_n), and transition (Mg_t) metals. The transition metal is in the same location as the previous ground-state metal-binding site. The localized relative charges are indicated.

basic side chain in A- and B-family DNA polymerases²⁸. The accumulating negative charge on $O_{P\alpha-\beta}$ also promotes protonation of PP_i , which has been proposed to be a key rate-limiting step²⁹. Following product formation this metal could transfer to the nearby product metal-binding site, while the catalytic metal rapidly dissociates owing to the loss of a coordinating ligand ($O3'$). Accordingly, the appearance/disappearance of cations around the active site represents an elegant ballet of electrons during DNA synthesis.

The structures captured here reveal how 8-oxo-dGTP escapes general polymerase discrimination checkpoints by modulating the highly charged DNA polymerase active site. Importantly, if 8-oxo-dGTP were to be inserted into a single-nucleotide gap, DNA ligase would be responsible for sealing the nick with a modified base pair. Abnormalities at the nick would hasten abortive ligation and stabilize the cytotoxic nick, thus increasing the probability for apoptotic cell signalling³⁰. Similarly, the extension efficiency during error-free DNA synthesis from the modified base pair would be reduced (that is, DNA synthesis would pause), promoting the generation of cytotoxic strand breaks. Therefore, the insertion and subsequent processing of 8-oxo-G in DNA offers a mechanism to manipulate the oxidative DNA damage response as well as target cancer cells that have an elevated metabolic rate.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 July; accepted 23 September 2014.

Published online 17 November 2014.

- Reuter, S., Gupta, S. C., Chaturvedi, M. M. & Aggarwal, B. B. Oxidative stress, inflammation, and cancer: how are they linked? *Free Radic. Biol. Med.* **49**, 1603–1616 (2010).
- Bont, R. D. & Larebeke, N. Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis* **19**, 169–185 (2004).
- Fraga, C. G., Shigenaga, M. K., Park, J. W., Dega, P. & Ames, B. N. Oxidative damage to DNA during aging: 8-hydroxy-2'-deoxyguanosine in rat organ DNA and urine. *Proc. Natl Acad. Sci. USA* **87**, 4533–4537 (1990).
- Total, M. D. & Baker, M. S. DNA precursor pool: a significant target for N-methyl-N-nitrosourea in C3H/10T1/2 clone 8 cells. *Proc. Natl Acad. Sci. USA* **79**, 2211–2215 (1982).
- Huber, K. V. *et al.* Stereospecific targeting of MTH1 by (S)-crizotinib as an anticancer strategy. *Nature* **508**, 222–227 (2014).
- Gad, H. *et al.* MTH1 inhibition eradicates cancer by preventing sanitation of the dNTP pool. *Nature* **508**, 215–221 (2014).
- Foti, J. J., Devadoss, B., Winkler, J. A., Collins, J. J. & Walker, G. C. Oxidation of the guanine nucleotide pool underlies cell death by bactericidal antibiotics. *Science* **336**, 315–319 (2012).
- Shibutani, S., Takeshita, M. & Grollman, A. P. Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* **349**, 431–434 (1991).
- Colussi, C. *et al.* The mammalian mismatch repair pathway removes DNA 8-oxodGMP incorporated from the oxidized dNTP pool. *Curr. Biol.* **12**, 912–918 (2002).
- Pursell, Z. F., McDonald, J. T., Mathews, C. K. & Kunkel, T. A. Trace amounts of 8-oxo-dGTP in mitochondrial dNTP pools reduce DNA polymerase γ replication fidelity. *Nucleic Acids Res.* **36**, 2174–2181 (2008).
- Oda, Y. *et al.* NMR studies of a DNA containing 8-hydroxydeoxyguanosine. *Nucleic Acids Res.* **19**, 1407–1412 (1991).
- Amouroux, R., Campalans, A., Epe, B. & Radicella, J. P. Oxidative stress triggers the preferential assembly of base excision repair complexes on open chromatin regions. *Nucleic Acids Res.* **38**, 2878–2890 (2010).
- Cabelof, D. C., Raffoul, J. J., Yanamadala, S., Guo, Z. & Heydari, A. R. Induction of DNA polymerase β -dependent base excision repair in response to oxidative stress *in vivo*. *Carcinogenesis* **23**, 1419–1425 (2002).
- Beard, W. A. & Wilson, S. H. Structure and mechanism of DNA polymerase β . *Biochemistry* **53**, 2768–2780 (2014).
- Donigan, K. A. *et al.* Human *POLB* gene is mutated in high percentage of colorectal tumors. *J. Biol. Chem.* **287**, 23830–23839 (2012).
- Beese, L. S. & Steitz, T. A. Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J.* **10**, 25–33 (1991).
- Sawaya, M. R., Prasad, P., Wilson, S. H., Kraut, J. & Pelletier, H. Crystal structures of human DNA polymerase β complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* **36**, 11205–11215 (1997).
- Batra, V. K. *et al.* Magnesium induced assembly of a complete DNA polymerase catalytic complex. *Structure* **14**, 757–766 (2006).
- Batra, V. K. *et al.* Mutagenic conformation of 8-oxo-7,8-dihydro-2'-dGTP in the confines of a DNA polymerase active site. *Nature Struct. Mol. Biol.* **17**, 889–890 (2010).
- Miller, H., Prasad, R., Wilson, S. H., Johnson, F. & Grollman, A. P. 8-OxodGTP incorporation by DNA polymerase β is modified by active-site residue Asn279. *Biochemistry* **39**, 1029–1033 (2000).
- Wang, Y. L. & Schlick, T. Distinct energetics and closing pathways for DNA polymerase β with 8-oxoG template and different incoming nucleotides. *BMC Struct. Biol.* **7**, 7 (2007).
- Freudenthal, B. D., Beard, W. A. & Wilson, S. H. DNA polymerase minor groove interactions modulate mutagenic bypass of a templating 8-oxoguanine lesion. *Nucleic Acids Res.* **41**, 1848–1858 (2013).
- Brown, J. A., Duym, W. W., Fowler, J. D. & Suo, Z. Single-turnover kinetic analysis of the mutagenic potential of 8-oxo-7,8-dihydro-2'-deoxyguanosine during gap-filling synthesis catalyzed by human DNA polymerases λ and β . *J. Mol. Biol.* **367**, 1258–1269 (2007).
- Eckenroth, B. E., Fleming, A. M., Sweasy, J. B., Burrows, C. J. & Doublié, S. Crystal structure of DNA polymerase β with DNA containing the base lesion spiroiminodihydantoin in a templating position. *Biochemistry* **53**, 2075–2077 (2014).
- Koag, M. C., Min, K. & Lee, S. Structural basis for promutagenicity of 8-halogenated guanine. *J. Biol. Chem.* **289**, 6289–6298 (2014).
- Freudenthal, B. D., Beard, W. A., Shock, D. D. & Wilson, S. H. Observing a DNA polymerase choose right from wrong. *Cell* **154**, 157–168 (2013).
- Nakamura, T., Zhao, Y., Yamagata, Y., Hua, Y.-j. & Yang, W. Watching DNA polymerase η make a phosphodiester bond. *Nature* **487**, 196–201 (2012).
- Beard, W. A. & Wilson, S. H. Structural insights into the origins of DNA polymerase fidelity. *Structure* **11**, 489–496 (2003).
- Sucato, C. A. *et al.* DNA polymerase β fidelity: halomethylene-modified leaving groups in pre-steady-state kinetic analysis reveal differences at the chemical transition state. *Biochemistry* **47**, 870–879 (2008).
- Harris, J. L. *et al.* Aprataxin, poly-ADP ribose polymerase 1 (PARP-1) and apurinic endonuclease 1 (APE1) function together to protect the genome against oxidative damage. *Hum. Mol. Genet.* **18**, 4102–4117 (2009).

Acknowledgements We thank the Collaborative Crystallography group at the National Institute of Environmental Health Sciences for help with data collection and analysis. We thank L. Pedersen for discussions. Use of the advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract W-31-109-Eng-38. This research was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences (project numbers Z01-ES050158 (to S.W.), Z01-ES050161 (to S.W.), and ZIC-ES043010 (to L.P.)) and in association with National Institutes of Health grant 1U19CA105010. We are grateful for computational support for the molecular dynamics simulations from the HPC clusters at NYU as well as the Blue Gene at CCNI. Support from Philip Morris USA Inc. and Philip Morris International to T.S. is gratefully acknowledged.

Author Contributions B.F., W.B., and S.W. designed the project. B.F. performed crystallography. D.S. did the kinetic analyses. T.K. and T.S. did the molecular dynamics simulations. L.P. did the quantum mechanical analysis. B.F., W.B., and S.W. prepared the manuscript. All authors discussed the results and commented on the manuscript.

Author Information Atomic coordinates and structure factors for the reported crystal structures have been deposited in the Protein Data Bank under accession numbers 4UAW, 4UAY, 4UAZ, 4UB1, 4UB2, 4UB3, 4UB4, 4UB5, 4UBB, and 4UBC. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.H.W. (wilson5@niehs.nih.gov).

METHODS

DNA sequences. To generate the 16-base oligonucleotide the following DNA sequences were used for crystallization (coding nucleotide is underlined): template, 5'-CCG ACA/C GCG CAT CAG C-3'; primer, 5'-GCT GAT GCG C-3'; downstream, 5'-GTC GG-3'. The downstream sequence was 5'-phosphorylated. The kinetic studies required extending the downstream and upstream sequences to employ a 34-base oligonucleotide DNA substrate. The sequence of the template strand was 5'-GTA CCC GGG GAT CCG TAC A/CGC GCA TCA GCT GCA G-3'. The underlined A/C represents the coding nucleotide as either an Ad or Cy. DNA substrates for single-nucleotide gap filling DNA synthesis measurements were prepared by annealing three purified oligonucleotides. Each oligonucleotide was suspended in 10 mM Tris-HCl, pH 7.4, and 1 mM EDTA and the concentration was determined from their ultraviolet absorbance at 260 nm. The annealing reactions were performed by incubating a solution of primer with downstream and template oligonucleotides (1:1.2:1.2 molar ratio, respectively) at 95 °C for 5 min, followed by 65 °C for 30 min, and finally cooling 1 °C min⁻¹ to 10 °C in a PCR thermocycler.

Protein expression, crystallization, and structural determination. Human wild-type, R283A, and R283K DNA polymerase β were overexpressed in *E. coli* and purified as described previously³¹. Binary complex crystals with a templating cytosine or adenine in a 1-nucleotide gapped DNA were grown as previously described¹⁸. The time-lapse crystallography was performed as previously described and is briefly summarized here²⁶. Binary pol β :DNA complex crystals were first transferred to a cryosolution containing 15% ethylene glycol, 50 mM imidazole, pH 7.5, 20% PEG3350, 90 mM sodium acetate, 3 mM 8-oxo-dGTP or dGTP, and 50 mM CaCl₂ for 1 h. These ground state ternary complex crystals were then transferred to a cryosolution containing 200 mM MgCl₂ or MnCl₂ for varying times. All reactions were stopped by freezing the crystals at 100 K before data collection at the home source, 1.54 Å, or the Advanced Photon Source, 1.0 Å (Argonne National Laboratory). In-house data collection was done on a SATURN92 CCD (charge-coupled device) detector system mounted on a MiraMax-007HF rotating anode generator at a wavelength of 1.54 Å. This allows for anomalous data detection after phasing by molecular replacement. Remote data collection was done at the Southeast Regional Collaborative Access Team BM-22 beamline at the Advanced Photon Source (Argonne National Laboratory) at a wavelength of 1.0 Å, with the MAR225 area detector. Data were processed and scaled using the HKL2000 software package³². Initial models were determined using molecular replacement with the open binary (Protein Data Bank accession number 3ISB) or closed ternary (Protein Data Bank accession number 2FMS) structures of pol β , and all R_{free} flags were taken from the starting model. Refinement used PHENIX and model building used Coot^{33,34}. The metal-ligand coordination restraints were generated by ReadySet (PHENIX) and not used until the final rounds of refinement. Partial catalysis models were generated with both the reactant and product species, and occupancy refinement was performed. The figures were prepared in PyMol and all density maps were generated after performing simulated annealing³⁵. Ramachandran analysis determined that 100% of non-glycine residues lie in allowed regions and at least 97% in favoured regions.

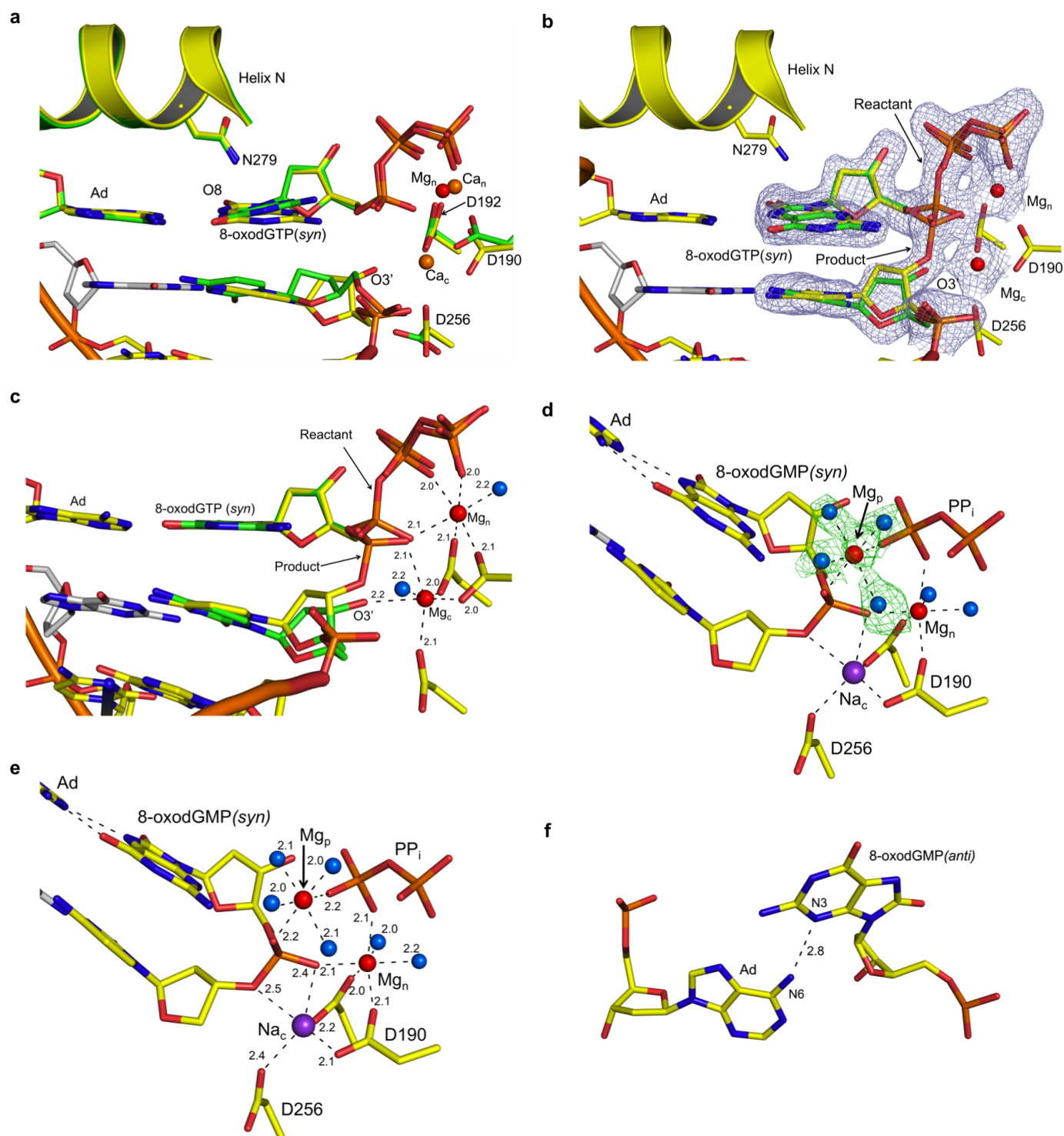
Kinetic characterization. Steady-state kinetic parameters for single-nucleotide gap filling reactions with wild-type enzyme were determined by initial velocity measurements as described previously³⁶. Unless noted otherwise, enzyme activities were determined using a standard reaction mixture containing 50 mM Tris-HCl, pH 7.4 (37 °C), 100 mM KCl, 10 mM MgCl₂, 1 mM dithiothreitol, 100 μ g ml⁻¹ bovine serum albumin, 10% glycerol, and 200 nM single-nucleotide gapped DNA. Enzyme concentrations and reaction time intervals were chosen so that substrate depletion or product inhibition did not influence initial velocity measurements. Owing to the low activity of the Arg 283 mutants (alanine and lysine), the catalytic efficiencies were determined by single-turnover analysis as described before except that the enzyme/DNA ratio was 10 (ref. 37). Reactions were quenched with 0.3 M EDTA and mixed with an equal volume of 95% formamide dye. The substrates and products were separated on 16% denaturing (8 M urea) polyacrylamide gels. Since a 6-carboxyfluorescein 5'-labelled primer was used in these assays, the substrates and products were quantified using a GE Typhoon 8600 phosphorimager in fluorescence mode (532 nm laser, 526 short-pass filter). Kinetic parameters were determined by fitting the rate data to a hyperbolic equation. When the observed rates could not be saturated owing to weak substrate binding, the data were fitted to an alternative form of the equation to extract catalytic efficiency (k_{cat}/K_M , best-fit initial slope); $k_{obs} = ((k_{cat}/K_M) \times [S]) / (1 + ([S]/K_M))$. The mean and standard error of at least two independent determinations are illustrated in plots that highlight substrate discrimination.

Quantum mechanical analysis. The three systems are shown in Extended Data Fig. 4 and described here. They are (1) 8-oxo-GTP(*anti*) with the Ca_c, Ca_n, Ca_g ions, the coordinating eight water molecules, and the three acetate ions (mimicking Asp 190, 192, and 256 of pol β); (2) all atoms in system 1 except for the Ca_g and its coordinating water molecules; (3) dGTP(*anti*) with the Ca_c, Ca_n, Ca_g ions, the coordinating

eight water molecules, and the three acetate ions (mimicking Asp 190, 192, and 256 of pol β). The base of 8-oxo-GTP was optimized under minimum constraints to keep the geometries closer to the crystallographic structure at the B3LYP level of theory with the 6-31+g* basis set using the program Gaussian09.D01 (ref. 38). All charges were calculated using the ChelpG procedure in Gaussian09.D01 at the 6-31+g** basis set level³⁹.

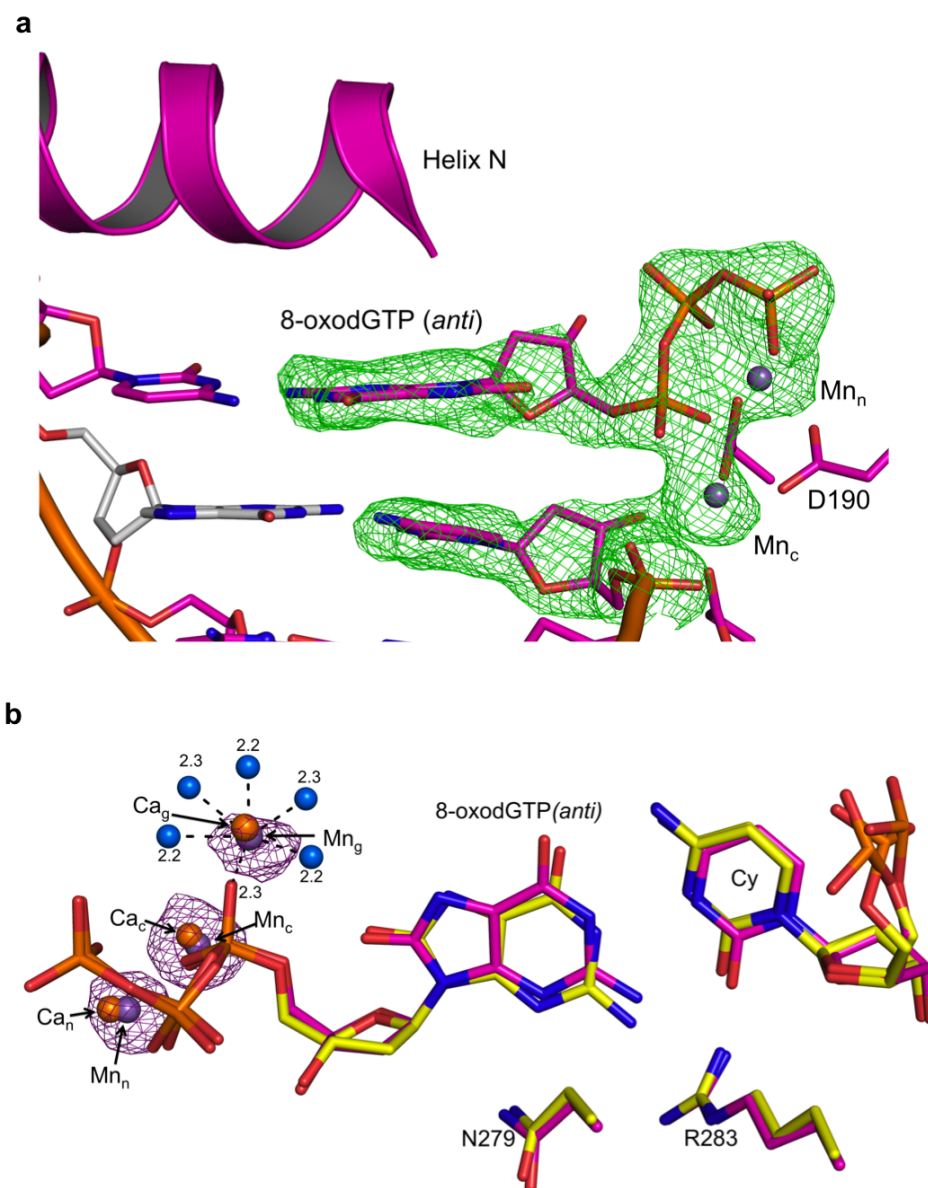
Molecular dynamics simulations. Two reference systems were simulated by molecular dynamics for 80 ns. System 1 was prepared using the entire ground state 8-oxo-dGTP(*anti*):Cy crystal structure (0 s). It consisted of the DNA, 8-oxo-dGTP(*anti*), pol β , 302 crystal water molecules, and three Mg²⁺ ions that replaced the Ca²⁺ ions in the crystal structure at the catalytic, nucleotide, and ground metal-binding sites. System 2 was prepared as system 1, but with 8-oxo-dGTP replaced by dGTP(*anti*). For all structures, missing atoms were added using CHARMM⁴⁰, and the disordered residues 1–10 of polymerase were attached by Accelrys Discovery Studio. Both systems were solvated with TIP3P water molecules. The smallest image distance between the solute and the faces of the periodic cubic cell was set to 12 Å. The total number of water atoms was 49,089 and total number of atoms 55,521. Neutralizing ions (Na⁺) and 150 mM NaCl were added to both systems. All of the Na⁺ and Cl⁻ ions were placed at least 8 Å away from each other, pol β , and the DNA. Both systems were minimized with fixed positions for all heavy atoms of pol β and DNA for 10,000 steps. The equilibration process was started with a 200 ps simulation at 300 K using Langevin dynamics, while keeping all the heavy atoms of pol β and DNA fixed. This was followed by unconstrained minimization consisting of 20,000 steps. The systems were then equilibrated for 500 ps at constant pressure and temperature. Pressure was maintained at 1 atmosphere using the Langevin piston method⁴¹. The temperature was maintained at 300 K using weakly coupled Langevin dynamics of non-hydrogen atoms with a damping coefficient of 10 ps⁻¹. The systems were simulated in periodic boundary conditions with full electrostatics computed using the particle mesh Ewald method⁴². Short-range non-bonded terms were evaluated at every step using a 12 Å cutoff for van der Waals interactions and a smooth switching function. The production simulations were performed for 80 ns with a 2 fs timestep. The minimization, equilibration, and production molecular dynamics simulations were performed by the NAMD simulation package⁴³ with the CHARMM27 all-atom force field^{44–46}. The force field parameters for 8-oxo-G were adopted from earlier works^{47,48}. The average distances and standard deviations were calculated using molecular dynamics trajectories in the 50–80 ns time range.

- Beard, W. A. & Wilson, S. H. Purification and domain-mapping of mammalian DNA polymerase β . *Methods Enzymol.* **262**, 98–107 (1995).
- Otinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
- The AxPyMOL Molecular Graphics Plugin for Microsoft PowerPoint, v.1.0 (Schrodinger, 2010).
- Beard, W. A., Shock, D. D. & Wilson, S. H. Influence of DNA structure on DNA polymerase β active site function: extension of mutagenic DNA intermediates. *J. Biol. Chem.* **279**, 31921–31929 (2004).
- Beard, W. A., Shock, D. D., Yang, X.-P., DeLauder, S. F. & Wilson, S. H. Loss of DNA polymerase β stacking interactions with templating purines, but not pyrimidines, alters catalytic efficiency and fidelity. *J. Biol. Chem.* **277**, 8235–8242 (2002).
- Frisch, M. J. et al. Gaussian 09, Revision D.01 (Gaussian, 2009).
- Chirlian, L. E. & Francl, M. M. Atomic charges derived from electrostatic potentials: a detailed study. *J. Comput. Chem.* **8**, 894–905 (1987).
- Brooks, B. R. et al. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217 (1983).
- Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* **103**, 4613–4621 (1995).
- Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an N -log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
- Phillips, J. C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
- Foloppe, N. & MacKerell, J. A. D. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.* **21**, 86–104 (2000).
- MacKerell, A. D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
- MacKerell, A. D. & Banavali, N. K. All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J. Comput. Chem.* **21**, 105–120 (2000).
- Wang, Y., Arora, K. & Schlick, T. Subtle but variable conformational rearrangements in the replication cycle of *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4) may accommodate lesion bypass. *Protein Sci.* **15**, 135–151 (2006).
- Pavelites, J. J., Gao, J., Bash, P. A. & MacKerell, A. D. A molecular mechanics force field for NAD⁺, NADH, and the pyrophosphate groups of nucleotides. *J. Comput. Chem.* **18**, 221–239 (1997).



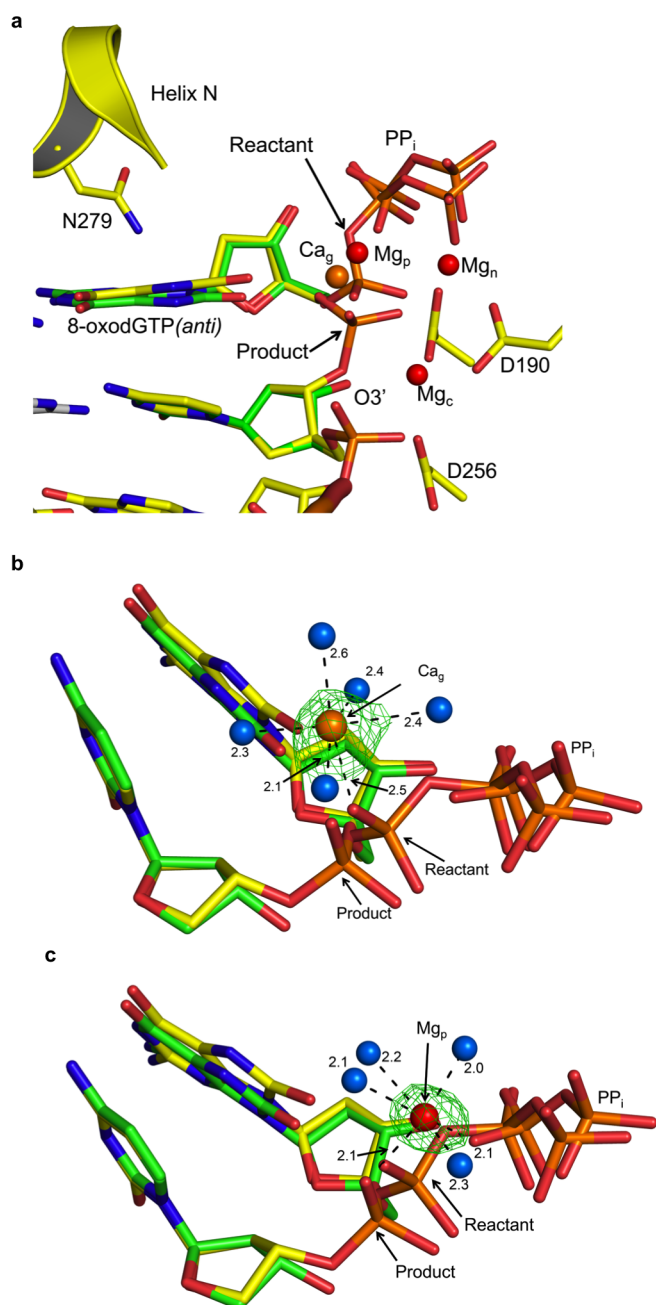
Extended Data Figure 1 | Mutagenic 8-oxo-dGTP insertion opposite adenine. **a**, Overlay of the ternary complex for 8-oxo-dGTP(*syn*):Ad generated with Ca^{2+} or a dideoxy-terminated primer (Protein Data Bank accession number 3MBY) is shown in yellow and green, respectively (root mean squared deviation of 0.17 Å). **b**, The pol β active site is shown with a $2F_o - F_c$ map contoured at 1.5σ after a 20 s soak. Key active site residues are indicated and Mg^{2+} ions are shown as red spheres. The reactant 8-oxo-dGTP and product 8-oxo-dGMP are shown in green and yellow respectively. **c**, A focused view

of **b** with the density removed. Coordinating waters (blue) and their distances (in ångströms) to active site metals are shown. **d**, The active site following a 40 s soak is shown with an omit map (3σ) for the Mg_p and coordinating waters. **e**, The coordination distances (in ångströms) for the Na_c , Mg_p , and Mg_n metals are indicated for the closed product complex after a 40 s soak. **f**, The 8-oxo-dGMP(*anti*):Ad contact between N3 and N6 of 8-oxo-dGMP and Ad respectively is shown for the open product complex after a 90 s soak.

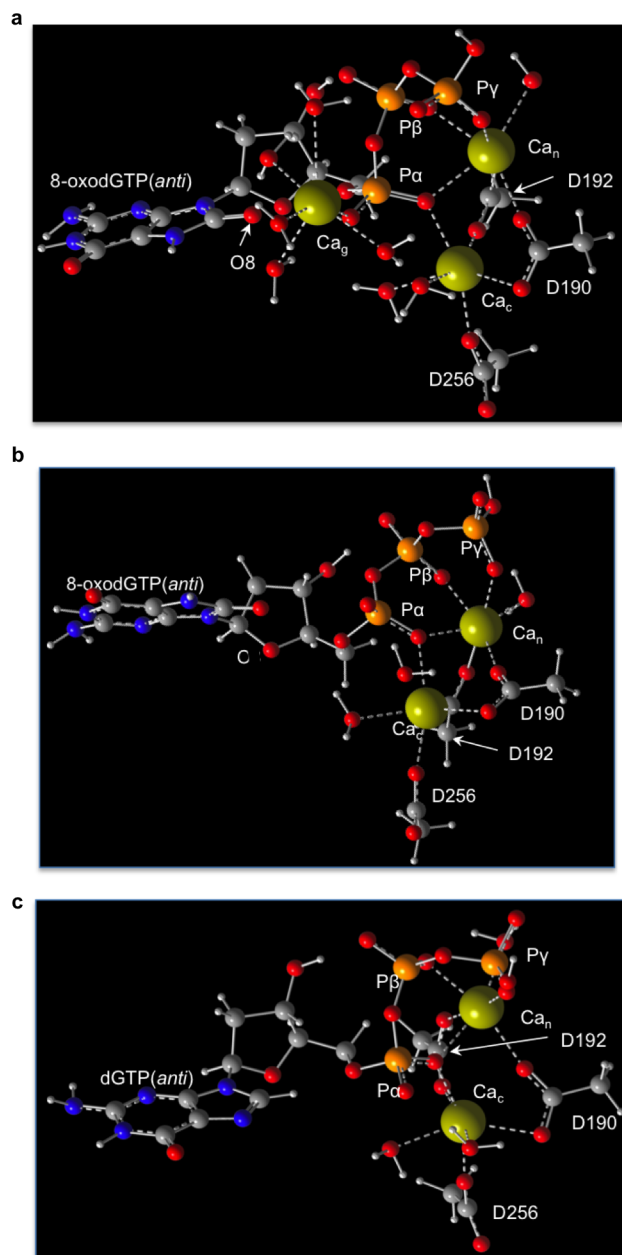


Extended Data Figure 2 | Pre-catalytic ground state with 8-oxo-dGTP and templating cytosine after a 5 s soak in MnCl₂. **a**, The pre-catalytic pol β active site is shown with an omit map (3σ). The ground state metal (Mn_g) has been removed for clarity. **b**, The view is a 90° rotation relative to **a**. An overlay of

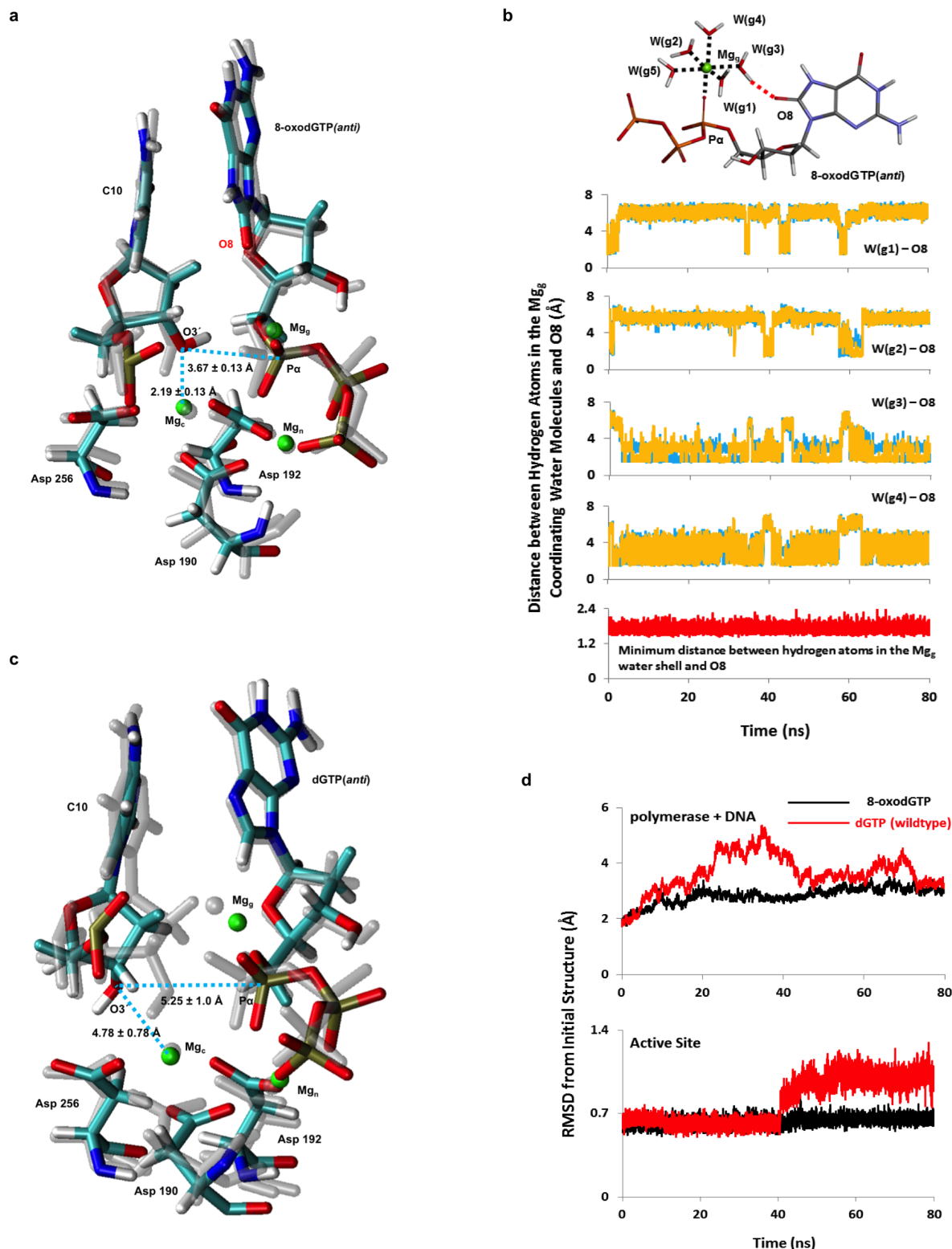
the 8-oxo-dGTP(*anti*) with Ca²⁺ and Mn²⁺ is shown in yellow and purple respectively. The anomalous density map contoured at 5σ for the Mn²⁺ ions is shown in purple. The Mn_g coordinating water molecules are shown in blue and the distances (in ångströms) are indicated.



Extended Data Figure 3 | Reaction with 8-oxo-dGTP opposite templating cytosine. **a**, Focused view of the active site following a 40 s soak is shown with key residues indicated; density has been removed for clarity (see Fig. 2e for density). **b**, An omit map (3σ) for Ca_g is shown. Coordinating waters are shown in blue (distances in ångströms). **c**, An omit map (3σ) for Mg_p is shown.



Extended Data Figure 4 | Quantum mechanical computational models for 8-oxo-dGTP(*anti*) and dGTP(*anti*). The models used for the quantum mechanical computational studies with the calcium ions, oxygen, phosphates, carbon, nitrogen, and protons shown in green, red, orange, grey, blue, and white, respectively. The key atoms and Asp 190, Asp 192, and Asp 256 mimics are indicated. **a**, The 8-oxo-dGTP(*anti*) with three calcium ions and eight water molecules. **b**, The 8-oxo-dGTP(*anti*) with two calcium ions and three water molecules. **c**, The dGTP(*anti*) with two calcium ions and three water molecules.



Extended Data Figure 5 | Molecular dynamics simulation analysis of 8-oxo-dGTP(*anti*) and dGTP(*anti*) opposite Cy. **a**, The 8-oxo-dGTP(*anti*) opposite Cy at 80 ns superimposed upon the initial structure. A multicolour code based on atom type is used for the final molecular dynamics structure, whereas the reference initial structure is shown in light grey. The catalytic (Mg_c), nucleotide (Mg_n), and ground (Mg_g) magnesium metal ions are shown in green, and average distances over the course of the simulation are indicated for $P\alpha$ -O3' and Mg_c -O3'. **b**, Distance distributions between hydrogen atoms in the water shell and O8 in the 8-oxo-dGTP(*anti*):Cy simulation. A snapshot of the 8-oxo-dGTP, Mg_g , and water shell (W(g1–g5)) is plotted at top. Black and red dotted lines indicate Mg_g coordination and a

hydrogen-bonding interaction between a water molecule and O8, respectively. Four of the five water molecules in the water shell (W(g1–g4)) contribute to hydrogen-bonding interactions with O8. Blue and orange lines indicate distances between hydrogen atoms in each water molecule and O8. The red line in the bottom plot indicates the minimum distance between hydrogen atoms in the water shell and O8. **c**, The dGTP(*anti*) opposite Cy at 80 ns superimposed upon the initial structure (grey). Distances and ion labelling are as for **a**. **d**, Root mean squared deviation of the evolving molecular dynamics structure for the entire polymerase/DNA complex (top) and for the active site only (bottom), with respect to the crystal structure.

Extended Data Table 1 | Data collection and refinement statistics of pol β 8-oxo-dGTP insertion opposite adenine

	0 s	20 s	40 s	90 s
	Ground state	Reactant state	Product state	Product state
Templating Base	Adenine	Adenine	Adenine	Adenine
Data Collection				
Wavelength	1.00	1.00	1.00	1.00
Space group	P2 ₁	P2 ₁	P2 ₁	P2 ₁
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	50.9,79.9,55.5	50.8,79.9,55.4	50.9,80.4,55.4	55.1,80.8,55.5
<i>a</i> , <i>b</i> , <i>c</i> (°)	90,107.6,90	90,107.6,90	90,107.7,90	90,109.6,90
Resolution (Å)	50.0–1.90	50.0–1.88	50.0–1.98	50.0–2.35
<i>R</i> _{sym} or <i>R</i> _{merge} (%)	6.8 (52.8)	5.5 (34.2)	6.0 (28.4)	5.2 (33.4)
<i>I</i> / <i>s</i>	28.9 (4.4)	23.3 (2.3)	24.2 (2.8)	24.3 (2.3)
Completeness (%)	98.0 (97.7)	95.6 (68.9)	96.6 (72.1)	97.3 (77.0)
Redundancy	5.7 (4.0)	3.3 (1.8)	3.5 (2.3)	3.5 (2.5)
Refinement				
Resolution (Å)	1.90	1.88	1.98	2.35
No. reflections	57079	56411	51562	34952
<i>R</i> _{work} / <i>R</i> _{free}	17.5/23.2	17.4/21.4	17.7/23.9	20.8/26.4
No. atoms				
Protein	2677	2673	2673	2593
DNA	659	681	681	681
Water	341	278	274	86
B-factors (Å ²)				
Protein	25.9	29.2	26.2	40.5
DNA/8oxo/PP _i	26.2/18.7/-	36.5/18.2/25.1	35.0/20.9/28.1	38.2/57/-
Water	27.8	33.1	31.57	30.4
R.m.s deviations				
Bond length (Å)	0.01	0.01	0.01	0.01
Bond angles (°)	1.20	1.00	1.09	1.16
Reaction Ratio				
Pol β conformation	closed	closed	closed	open
Ratio of RS/PS [†]	1.0/0	0.4/0.6	0/1.0	0/1.0
Occupancy				
Metal C/N/G/P [‡]	1.0/1.0/-/-	1.0/1.0/-/-	-/1.0/-/0.7	-
PP _i	-	0.6	1.0	-
PDB ID	4UAW	4UAZ	4UAY	4UB1

* Highest resolution shell is shown in parentheses.

† RS and PS represent the reactant state and product states, respectively.

‡ Metal C/N/G/P refers to the catalytic, nucleotide, ground, and product metal-binding sites, respectively.

Extended Data Table 2 | Key DNA and structural parameters near the pol β active site*

Complex	Sugar pucker		Buckle (NBP) [‡]	Shear (NBP)
	Primer [†]	Incoming		
dC–dGTP	C3'-endo	C3'-endo	8.5	-0.17
dC–8oxodGTP	C3'-endo	C4'-exo	-13.1	0.43
dA–8oxodGTP	C3'-endo	C3'-endo	0.7	0.03
dG–dATP [§]	C2'-endo	C3'-endo	43.2	0.61
dA–8oxodGTP	C2'-endo	C4'-exo	-13.9	-0.06

* Parameters extracted with 3DNA.

† Primer terminus.

‡ Nascent base pair.

§ Protein Data Bank accession number 4LVS.

|| Protein Data Bank accession number 3MBY.

Extended Data Table 3 | Data collection and refinement statistics of pol β insertion opposite cytosine with 8-oxo-dGTP and dGTP

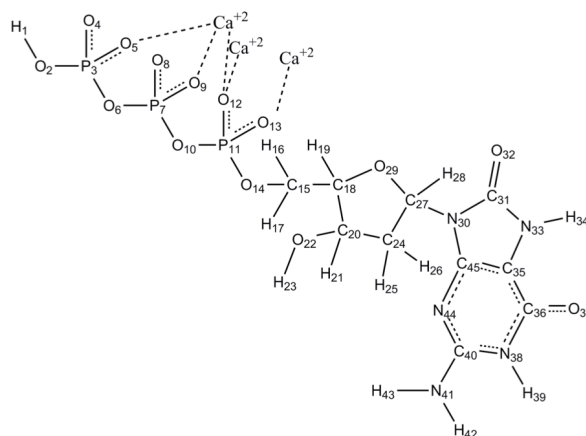
	0 s	40 s	60 s	120 s	5 s (MnCl ₂)	0 s (dGTP)
	Ground state	Reactant state	Product State	Product state	Ground state	Ground state
Templating Base	Cytosine	Cytosine	Cytosine	Cytosine	Cytosine	Cytosine
Data Collection						
Wavelength	1.54	1.00	1.54	1.54	1.54	1.54
Space group	P2 ₁	P2 ₁	P2 ₁	P2 ₁	P2 ₁	P2 ₁
Cell dimensions						
<i>a</i> , <i>b</i> , <i>c</i> (Å)	50.7,79.8,55.5	50.8,79.8,55.6	50.7,80.1,55.4	55.1,79.6,55.7	55.1,77.7,55.1	50.6,79.4,55.5
<i>a</i> , <i>b</i> , <i>g</i> (°)	90,107.5,90	90,107.6,90	90,107.8,90	90,109.7,90	90,114.2,90	90,107.5,90
Resolution (Å)	50.0-1.96	50.0-1.85	50.0-2.06	50.0-2.51	50.0-2.15	50.0-1.96
<i>R</i> _{sym} or <i>R</i> _{merge} (%)	6.0 (42.1)	5.4 (27.5)	8.3 (51.1)	7.6 (62.3)	6.9 (51.6)	6.8 (51.3)
<i>I</i> / <i>σ</i>	20.4 (2.1)	20.9 (2.9)	16.3 (2.1)	17.6 (2.2)	13.3 (2.3)	19.6 (2.2)
Completeness (%)	98.4 (87.7)	98.1 (79.0)	97.4 (91.2)	99.5 (99.7)	99.6 (99.5)	99.5 (95.8)
Redundancy	3.0 (1.7)	3.2 (1.9)	4.7 (3.3)	3.6 (3.5)	3.2 (2.9)	4.8 (2.3)
Refinement						
Resolution (Å)	2.0	1.90	2.06	2.51	2.15	1.96
No. reflections	51032	63744	21786	25634	39111	49693
<i>R</i> _{work} / <i>R</i> _{free}	18.2/25.4	17.6/22.3	19.1/25.2	20.5/27.5	21.2/27.4	17.9/22.8
No. atoms						
Protein	2673	2673	2673	2593	2673	2674
DNA	627	661	661	661	627	627
Water	304	293	172	63	101	235
B-factors (Å ²)						
Protein	27.7	28.8	32.3	38.9	42.1	36.1
DNA/8oxo/PP _i	36.6/18.8/-	38.0/21.0/20.5	41.4/25.8/28.1	34.8/65.4/-	44.1/38.3/-	42.6/25.0/-
Water	29.4	39.5	30.0	20.0	31.4	32.8
R.m.s deviations						
Bond length (Å)	0.01	0.01	0.006	0.01	0.007	0.009
Bond angles (°)	1.17	1.08	1.20	1.26	1.1	1.17
Reaction Ratio						
Pol β conformation	closed	closed	closed	open	closed	closed
Ratio of RS/PS [†]	1.0/0	0.6/0.4	0/1.0	0/1.0	1.0/0	1.0/0
Occupancy						
Metal C/N/G/P [‡]	1.0/1.0/1.0/-	1.0/1.0/0.6/0.4	-/1.0/-/0.6	-	1.0/1.0/1.0/-	1.0/1.0/-/-
PP _i	-	0.4	1.0	-	-	-
PDB ID	4UBC	4UBB	4UB3	4UB2	4UB5	4UB4

* Highest resolution shell is shown in parentheses.

[†] RS and PS represent the reactant state and product states, respectively.[‡] Metal C/N/G/P refers to the catalytic, nucleotide, ground, and product metal-binding sites, respectively.

Extended Data Table 4 | Quantum mechanical point charges for each atom of 8-oxo-dGTP(*anti*) and dGTP(*anti*) with either two or three calcium ions

Number*	Atom	8-oxodGTP 3 metals	8-oxodGTP 2 metals	dGTP 2 metals	8-oxodGTP (MD)
1	H	0.41	0.39	0.4	N/A
2	O	-0.67	-0.74	-0.68	-0.9
3	P	1.49	1.4	1.42	1.10
4	O	-0.9	-0.82	-0.87	-0.9
5	O	-0.83	-0.84	-0.82	-0.9
6	O	-0.62	-0.56	-0.58	-0.86
7	P	1.49	1.33	1.32	1.50
8	O	-0.81	-0.83	-0.79	-0.82
9	O	-0.88	-0.89	-0.82	-0.82
10	O	-0.75	-0.41	-0.51	-0.74
11	P	1.44	1.32	1.33	1.5
12	O	-0.91	-0.86	-0.91	-0.82
13	O	-0.97	-0.89	-0.77	-0.82
14	O	-0.22	-0.17	-0.33	-0.62
15	C	0.03	0.01	0	-0.08
16	H	0.11	0.06	0.14	0.09
17	H	0.05	0.04	0.05	0.09
18	C	0.14	0.25	0.13	0.16
19	H	0.03	-0.04	0.02	0.09
20	C	0.63	0.68	0.72	0.14
21	H	-0.12	-0.08	-0.14	0.09
22	O	-0.82	-0.88	-0.85	-0.66
23	H	0.48	0.49	0.49	0.43
24	C	-0.16	-0.16	-0.20	-0.18
25	H	0.05	0.01	0.01	0.09
26	H	-0.01	-0.02	0.04	0.09
27	C	0.59	0.60	0.44	0.16
28	H	-0.05	-0.08	0.02	0.09
29	O	-0.55	-0.58	-0.55	-0.5
30	N	-0.24	-0.26	-0.14	0.16
31	C	0.66	0.71	0.33	0.41
32†	O/H	-0.63	-0.61	0.08	-0.64
33	N	-0.51	-0.54	-0.57	-0.34
34‡	H	0.37	0.38	-	0.38
35	C	-0.19	-0.18	-0.02	0
36	C	0.77	0.77	0.78	0.33
37	O	-0.62	-0.66	-0.63	-0.58
38	N	-0.81	-0.83	-0.88	-0.23
39	H	0.42	0.41	0.42	0.32
40	C	0.91	0.92	0.97	0.57
41	N	-0.86	-0.89	-0.88	-0.87
42	H	0.38	0.37	0.36	0.41
43	H	0.36	0.35	0.33	0.42
44	N	-0.72	-0.72	-0.78	-0.57
45	C	0.44	0.43	0.39	0.23



The molecular dynamics point charges used with three magnesium ions are shown for reference in the last column. The units are in electron charge (e^-) and the key is shown for reference.

* The position of each atom is shown in the chemical structure cartoon below the table.

† The oxygen or proton corresponds to 8-oxo-dGTP and dGTP respectively.

‡ Proton corresponds to 8-oxo-dGTP only.

Extended Data Table 5 | Average distances in the active sites in 8-oxo-dGTP(*anti*) and dGTP(*anti*)

Distance	8-oxodGTP (Å)	dGTP (Å)
Mg _c – OD2 (ASP190)	1.81 ± 0.04	1.81 ± 0.04
Mg _c –OD1 (ASP192)	1.80 ± 0.04	1.81 ± 0.04
Mg _c – OD2 (ASP256)	1.80 ± 0.04	1.81 ± 0.04
Mg _c – O1A (Pα)	3.18 ± 0.13	3.34 ± 0.25
Mg _c – O3' (C10)	2.19 ± 0.13	4.78 ± 0.42
Mg _c – W(c)	1.95 ± 0.05	2.00 ± 0.07
Mg _c – W(c)*	N/A	1.98 ± 0.07
Pα – O3' (C10)	3.67 ± 0.13	5.25 ± 1.01
Mg _n – OD1 (ASP190)	1.86 ± 0.05	1.86 ± 0.05
Mg _n – OD2 (ASP192)	1.87 ± 0.05	1.88 ± 0.05
Mg _n – O1A (Pα)	1.90 ± 0.06	1.92 ± 0.06
Mg _n – O2B (Pβ)	1.90 ± 0.06	1.91 ± 0.06
Mg _n – O3G (Pγ)	1.85 ± 0.05	1.84 ± 0.05
Mg _n – W(n)	2.00 ± 0.06	2.03 ± 0.07
Mg _g – O2A (Pα)	1.82 ± 0.04	1.82 ± 0.04
Mg _g – W(g1)	1.98 ± 0.06	1.99 ± 0.06
Mg _g – W(g2)	1.98 ± 0.06	1.99 ± 0.06
Mg _g – W(g3)	2.00 ± 0.07	1.99 ± 0.06
Mg _g – W(g4)	2.00 ± 0.07	2.01 ± 0.07
Mg _g – W(g5)	1.99 ± 0.07	1.98 ± 0.06
W(g) – O8 (8-oxodGTP)	1.73 ± 0.11	N/A
W(g) – N7 (dGTP)	N/A	3.9 ± 0.34 Å

W(c), W(n), W(g1–g5) are water molecules bound to the Mg_c, Mg_n and Mg_g, respectively. The Mg_c in the dGTP(*anti*) system establishes a new coordination with a water molecular (W(c)*) after 40 ns. The average values are calculated using 50–80 ns range of molecular dynamics trajectories.

Structural insight into autoinhibition and histone H3-induced activation of DNMT3A

Xue Guo^{1,2*}, Ling Wang^{1,2*}, Jie Li¹, Zhanyu Ding³, Jianxiong Xiao¹, Xiaotong Yin¹, Shuang He¹, Pan Shi^{4,5,6}, Liping Dong^{7,8}, Guohong Li⁷, Changlin Tian^{4,5,6}, Jiawei Wang⁹, Yao Cong³ & Yanhui Xu^{1,2}

DNA methylation is an important epigenetic modification that is essential for various developmental processes through regulating gene expression, genomic imprinting, and epigenetic inheritance^{1–5}. Mammalian genomic DNA methylation is established during embryogenesis by *de novo* DNA methyltransferases, DNMT3A and DNMT3B^{6–8}, and the methylation patterns vary with developmental stages and cell types^{9–12}. DNA methyltransferase 3-like protein (DNMT3L) is a catalytically inactive paralogue of DNMT3 enzymes, which stimulates the enzymatic activity of Dnmt3a¹³. Recent studies have established a connection between DNA methylation and histone modifications, and revealed a histone-guided mechanism for the establishment of DNA methylation¹⁴. The ATRX–DNMT3–DNMT3L (ADD) domain of Dnmt3a recognizes unmethylated histone H3 (H3K4me0)^{15–17}. The histone H3 tail stimulates the enzymatic activity of Dnmt3a *in vitro*^{17,18}, whereas the molecular mechanism remains elusive. Here we show that DNMT3A exists in an autoinhibitory form and that the histone H3 tail stimulates its activity in a DNMT3L-independent manner. We determine the crystal structures of DNMT3A–DNMT3L (autoinhibitory form) and DNMT3A–DNMT3L–H3 (active form) complexes at 3.82 and 2.90 Å resolution, respectively. Structural and biochemical analyses indicate that the ADD domain of DNMT3A interacts with and inhibits enzymatic activity of the catalytic domain (CD) through blocking its DNA-binding affinity. Histone H3 (but not H3K4me3) disrupts ADD–CD interaction, induces a large movement of the ADD domain, and thus releases the autoinhibition of DNMT3A. The finding adds another layer of regulation of DNA methylation to ensure that the enzyme is mainly activated at proper targeting loci when unmethylated H3K4 is present, and strongly supports a negative correlation between H3K4me3 and DNA methylation across the mammalian genome^{9,10,19,20}. Our study provides a new insight into an unexpected autoinhibition and histone H3-induced activation of the *de novo* DNA methyltransferase after its initial genomic positioning.

To investigate how DNMT3A activity is regulated, we performed an *in vitro* DNA methylation assay using recombinant DNMT3A2 (residues 224–912), a catalytically active variant of DNMT3A consisting of Pro-Trp-Trp-Pro (PWWP), ADD, and CD domains (Fig. 1a)²¹. H3K4me0 (but not H3K4me3) peptide significantly stimulated the enzymatic activities of DNMT3A2 or DNMT3A2 in complex with the CD-like domain of DNMT3L (designated C^{DNMT3L}) (Fig. 1b and Extended Data Fig. 1a, b). A similar effect was observed for DNMT3A2 or DNMT3A2–C^{DNMT3L} using recombinant poly-nucleosomes carrying unmethylated histone H3 or H3K4me3 mimic (H3K_C4me3) (Fig. 1b and Extended Data Fig. 1c). Thus, the enzymatic activity of DNMT3A2 is stimulated by histone H3 (but not H3K4me3) tail either in the form of free peptide, or within nucleosome, in a DNMT3L-independent manner.

We next characterized the critical regions for histone H3-induced activation of DNMT3A. H3K4me0 peptide activated DNMT3A2 and ADD–CD (residues 476–912) proteins with comparable (approximately sixfold) activity enhancement, but did not stimulate activity of the CD domain (residues 627–912) in the absence or presence of C^{DNMT3L} (Extended Data Fig. 1d, e). ADD–CD protein with histone H3 peptide (residues 1–20) fused at the amino (N) terminus (H3–ADD–CD) showed significantly higher activity than ADD–CD, and could not be further activated by H3K4me0 peptide (Extended Data Fig. 1e). These results collectively indicate that the ADD domain (but not the PWWP domain) is required for histone H3-induced activation of DNMT3A.

Notably, the CD domain or CD–C^{DNMT3L} complex showed comparable activity to that of ADD–CD or ADD–CD–C^{DNMT3L} activated by histone H3 peptide (Extended Data Fig. 1d, e), suggesting that the ADD domain inhibit the activity of the CD domain and histone H3 release this inhibition. In support of this hypothesis, addition of ADD–linker (residues 476–626, Fig. 1a), but not H3–ADD–linker fusion protein inhibited the activity of the CD domain (Fig. 1c). These results confirm the existence of autoinhibition of DNMT3A, in which ADD–linker directly inhibits the enzymatic activity of the CD domain, and that this inhibition is released by histone H3 tail.

We determined the crystal structure of ADD–CD of DNMT3A in complex with C^{DNMT3L} (ADD–CD–C^{DNMT3L}) at 3.82 Å resolution (Fig. 1d and Extended Data Table 1). Although only one ADD–CD–C^{DNMT3L} complex was observed within an asymmetric unit of the crystals, gel-filtration analysis indicated a tetramer formation in solution. Two ADD–CD–C^{DNMT3L} complexes form dimer of dimers via the CD–CD interaction in a two-fold crystallographic symmetry (Extended Data Fig. 2). The complex structure adopts an ‘X’ shape with two CD domains located in the centre and the ADD and C^{DNMT3L} domains located in the four corners. In the ADD–CD–C^{DNMT3L} structure, CD–C^{DNMT3L} adopts a similar fold to that of mouse CD^{Dnmt3a}–C^{Dnmt3L} structure (Extended Data Fig. 3a)²².

The ADD and CD domains fold into two individual structural modules connected by the linker (Fig. 1d). The linker forms a twisted helix followed by an extended loop, which packs against a hydrophobic surface of the CD domain (Extended Data Fig. 3b, c). A loop region (residues 526–533) extends out of the ADD domain and inserts into a pocket in the CD domain (Fig. 1e and Extended Data Fig. 3d). This pocket is mainly formed by basic residues R790, R792, H789, and R831 of the CD domain. Three acidic residues (D529/D530/D531) and hydrophobic residues (Y526/Y528/Y533) of the ADD domain are brought into close proximity to the pocket. All residues that are involved in the intramolecular interaction are highly conserved in DNMT3A/3B, suggesting a conserved ADD–CD interface among DNMT3A/B sub-family

¹Fudan University Shanghai Cancer Center, Institute of Biomedical Sciences, Shanghai Medical College of Fudan University, Shanghai 200032, China. ²State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai 200433, China. ³National Center for Protein Science Shanghai, State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China. ⁴High Magnetic Field Laboratory, Chinese Academy of Sciences, Hefei 230031, China. ⁵National Laboratory for Physical Science at the Microscale, University of Science and Technology of China, Hefei 230026, China. ⁶School of Life Sciences, University of Science and Technology of China, Hefei 230026, China. ⁷National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Science, Beijing 100101, China. ⁸University of Chinese Academy of Science, Beijing 100049, China. ⁹State Key Laboratory of Biomembrane and Membrane Biotechnology, School of Life Sciences, Tsinghua University, Beijing 100084, China.

*These authors contributed equally to this work.

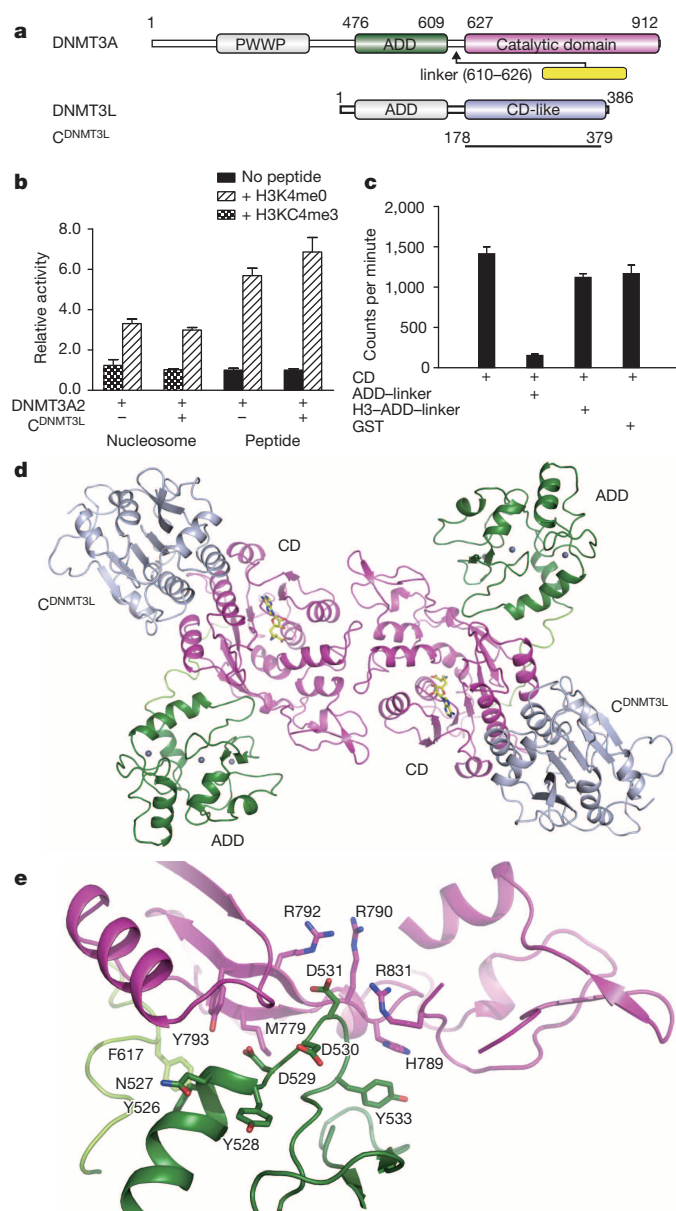


Figure 1 | Structure of the DNMT3A–DNMT3L complex in autoinhibitory form. **a**, Colour-coded domain architecture of human DNMT3A and DNMT3L. **b**, Relative activities of DNMT3A2 using peptides or reconstituted nucleosomes as substrates in the absence or presence of C^{DNMT3L}. **c**, Activities of the CD domain in the absence or presence of ADD–linker or H3–ADD–linker fusion proteins. **d**, Ribbon representations of the overall structure of the ADD–CD–C^{DNMT3L} complex. AdoHcy are shown in stick representation; zinc cations are shown as grey balls. The colour scheme is the same as in **a** and used in all structural figures. **e**, Close-up view of the ADD–CD interface. Critical residues for the interactions are shown in stick representation. Relative activities were calculated according to basal activity for each protein in the absence of H3 peptide. Error bars, s.d. for triplicate experiments.

members, but not DNMT3L (Extended Data Fig. 4). Structural comparison also indicated that ADD^{DNMT3L} would have steric hindrance with C^{DNMT3L} if DNMT3L adopts a similar conformation to that of DNMT3A (Extended Data Fig. 3e)¹⁵.

The intramolecular interaction was verified in the *in vitro* glutathione S-transferase (GST) pull-down assays, in which ADD–linker bound to the CD domain (Fig. 2a and Extended Data Fig. 5a–e). Mutations D529A, D531A, D529A/D531A, and Y526A/Y528A of ADD–linker, decreased the binding affinity to the CD domain, supporting their important role in mediating the ADD–CD interaction. In contrast, mutations R556A,

M548W, D530A, and Y533A of ADD–linker showed little effect on ADD–CD interaction (Fig. 2a and Extended Data Fig. 5a, b). Mutation R790A/R792A of the CD domain also impaired its binding affinity to the ADD–linker protein (Extended Data Fig. 5c). H3K4me0 (but not H3K4me3) peptide hampered ADD–CD interaction, suggesting that histone H3 tail releases autoinhibition of DNMT3A through disrupting this intramolecular interaction (Fig. 2b and Extended Data Fig. 5d).

We next mapped the critical regions for the autoinhibition of DNMT3A in the presence of C^{DNMT3L} (Fig. 2c). Compared with ADD–CD (Del-N476), Del-N522 showed a slight increase in enzymatic activity, whereas Del-N550, Del-N587, and Del-N610 significantly increased their activities (Fig. 2d). Although ADD and ADD–linker showed comparable binding affinity to the CD domain (Extended Data Fig. 5e), only ADD–linker (ADD–626) maintained the inhibitory function, while all other truncated proteins markedly decreased their inhibition on the activity of the CD domain (Fig. 2e). Replacement of residues 621–632 by a 12-residue Gly and Ser (GS) linker in ADD–CD partly released autoinhibition and showed less than twofold activity enhancement upon histone H3 stimulation (Extended Data Fig. 5f). The results indicate that the ADD domain and linker are both important for the autoinhibition of DNMT3A.

We next tested whether residues on the ADD–CD interface are important for the autoinhibition of DNMT3A. Compared with wild-type protein, mutations D529A, D531A, and Y526A/Y528A of the ADD–linker could barely inhibit the activity of the CD domain (Fig. 2f and Extended Data Table 2). Mutations D529A, D531A, and Y526A/Y528A of DNMT3A2 showed enzymatic activities that are comparable to or slightly lower than that of corresponding mutants activated by histone H3 peptide (Fig. 2g). A more obvious release of autoinhibition was observed in assays using recombinant poly-nucleosomes as substrate (Fig. 2h). As a negative control, mutation M548W of DNMT3A2 remained in an inhibitory form and was not activated by histone H3 tail because the mutant could not bind to histone H3 (ref. 18). Taken together, the ADD domain and linker function together to inhibit the activity of the CD domain, and ADD–CD association is essential for ADD-mediated autoinhibition of DNMT3A.

HhaI is a well-characterized bacterial DNA (cytosine-5) methyltransferase²³. Structural comparison of ADD–CD–C^{DNMT3L} and HhaI–DNA (Protein Data Bank (PDB) accession number 1MHT) complexes indicates that their catalytic domains adopt similar fold, and suggests that loops L1 and L2 of DNMT3A are potential regions for DNA interaction (Fig. 3a–c and Extended Data Fig. 4). Consistently, mutating basic residues K831, R836/N838, K841, and K844 on loop L2 impaired the enzymatic activity of DNMT3A (Fig. 3c and Extended Data Fig. 6a)²⁴. In the ADD–CD–C^{DNMT3L} structure, the ADD domain is located close to the loop L2 and may thus generate steric hindrance for DNA interaction (Fig. 3c).

In support of the analysis above, ADD–CD markedly decreased, but H3–ADD–CD showed comparable DNA-binding affinity to that of the CD domain (Fig. 3d). Histone H3 peptide partly restored DNA-binding affinity of ADD–CD or DNMT3A2, which may have resulted from relative weak interaction between histone H3 peptide and DNMT3A2 (Fig. 3d, e and Extended Data Fig. 6b). Consistent with their effect on the release of autoinhibition, mutations D529A, D531A, and Y526A/Y528A of DNMT3A2–C^{DNMT3L} partly rescued the DNA interaction (Fig. 3e), and histone H3 (but not H3K4me3) peptide increased the complex formation of ADD–CD–DNA, but not CD–DNA (Fig. 3f). Similar results were observed in electrophoretic mobility-shift assays (Extended Data Fig. 6c). Taken together, the ADD domain inhibits the activity of the CD domain by decreasing its DNA-binding affinity. This inhibition is restored by fusion of the histone H3 tail at the N terminus of ADD–CD, and is partly restored by addition of histone H3 peptide or mutations of residues on the ADD–CD interface.

We also determined the crystal structure of ADD–CD–C^{DNMT3L} in complex with histone H3 peptide (residues 1–12) at 2.90 Å resolution (Extended Data Table 1). The complex structure adopts a butterfly shape, with the ADD and C^{DNMT3L} domains resembling the wings (Fig. 4a). In

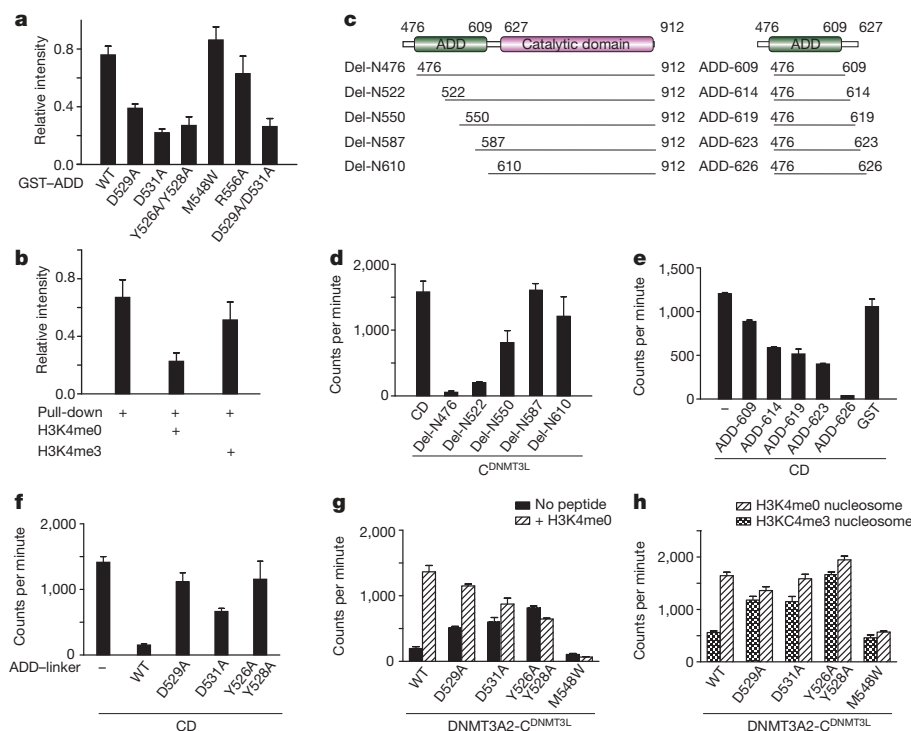


Figure 2 | ADD-CD interactions are important for autoinhibition of DNMT3A. **a**, GST pull-down assays for the ADD-CD interactions. Recombinant CD (residues 627–912) proteins were incubated with wild-type or mutant GST-ADD-linker proteins immobilized on glutathione resin. The bound proteins were analysed using SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie blue staining (Extended Data Fig. 5a). The assays were quantified by band densitometry. Error bars, s.d. for triplicate experiments. **b**, GST pull-down assays in the absence or presence of histone H3 peptide (H3K4me0 or H3K4me3) as indicated in Extended Data Fig. 5d. **c**, Schematic representation of DNMT3A proteins used for *in vitro* methyltransferase activity assays. **d**, Effect of various N-terminal deletions of ADD-CD on its enzymatic activity. **e**, Activities of the CD domain measured in the presence of various carboxy (C)-terminal deletions of ADD-linker proteins. **f**, Activities of the CD domain measured in the presence of wild-type and mutant ADD-linker proteins. **g**, **h**, Activities of wild-type or mutant DNMT3A2 measured using naked DNA (**g**) or poly-nucleosomes (**h**) as substrate. Error bars, s.d. for triplicate experiments.

the complex structure, two ADD-CD-C^{DNMT3L}-H3 complexes form a dimer of dimers through CD-CD interface in a pseudo-two-fold symmetry. CD-CD^{DNMT3L} adopts a similar fold to that of ADD-CD-C^{DNMT3L} (autoinhibitory form) and mouse CD^{DNMT3A}-C^{DNMT3L} structures, whereas the position of the ADD domain is obviously different in the structures of DNMT3A in autoinhibitory and active forms (Fig. 4a, b and Extended Data Fig. 7a, b).

In the ADD-CD-C^{DNMT3L}-H3 structure, the ADD-CD interaction is mediated by a network of hydrogen bonds and hydrophobic interactions (Extended Data Fig. 7c, d). A similar domain organization was observed in the DNMT3L-H3 structure¹⁵, suggesting a conserved ADD-CD

interface among DNMT3 family members when they adopt the active form (Extended Data Fig. 7e). The histone H3 peptide binds to the ADD domain by forming a three-stranded anti-parallel β -sheet with two β -strands of the ADD domain (Extended Data Fig. 7f). The side chain of H3K4 is stabilized by residues D529 and D531 of the ADD domain, and tri-methylation of H3K4 will disrupt the interactions. Mutations D529A, D531A, and M548W of ADD-CD abolished their binding affinity to histone H3 peptide (Extended Data Fig. 7g).

Comparison of the structures of ADD-CD-C^{DNMT3L} in autoinhibitory and active forms indicates two separate surfaces on the CD domain for ADD-CD interaction, and suggests a conformational change of DNMT3A

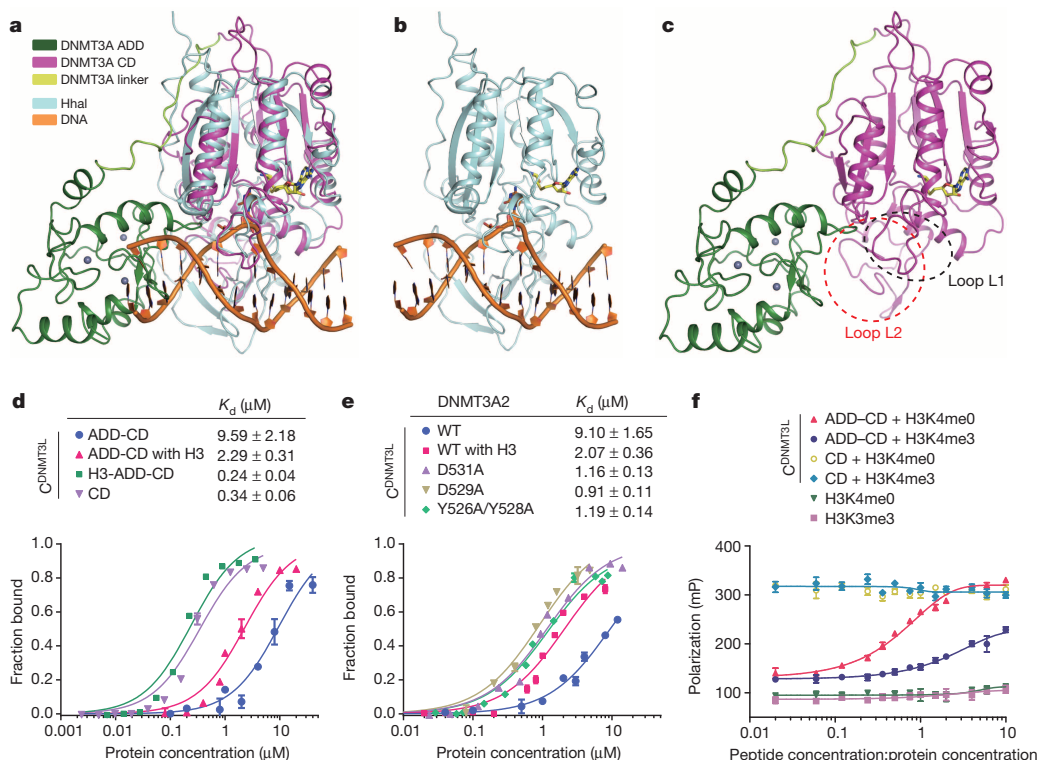


Figure 3 | The ADD domain inhibits DNA-binding affinity of the catalytic domain.

a, Superimposition of ADD-CD-C^{DNMT3L} and HhaI-DNA complex structures shown in ribbon representations with the C^{DNMT3L} domain omitted for simplicity. The colour scheme for the comparison is indicated. **b**, **c**, HhaI-DNA (**b**) and ADD-CD (**c**) structures are shown as in **a**. **d**, **e**, Superimposed fluorescence polarization plots for truncations (**d**) or mutants (**e**) of DNMT3A in the absence or presence of histone H3 peptide. **f**, Superimposed fluorescence polarization plots for protein-DNA complex formation of various truncations of DNMT3A in the absence or presence of histone H3 peptides. The levels of protein-DNA complex formation are represented by fluorescence polarization response (in milli-polarization units, mP).

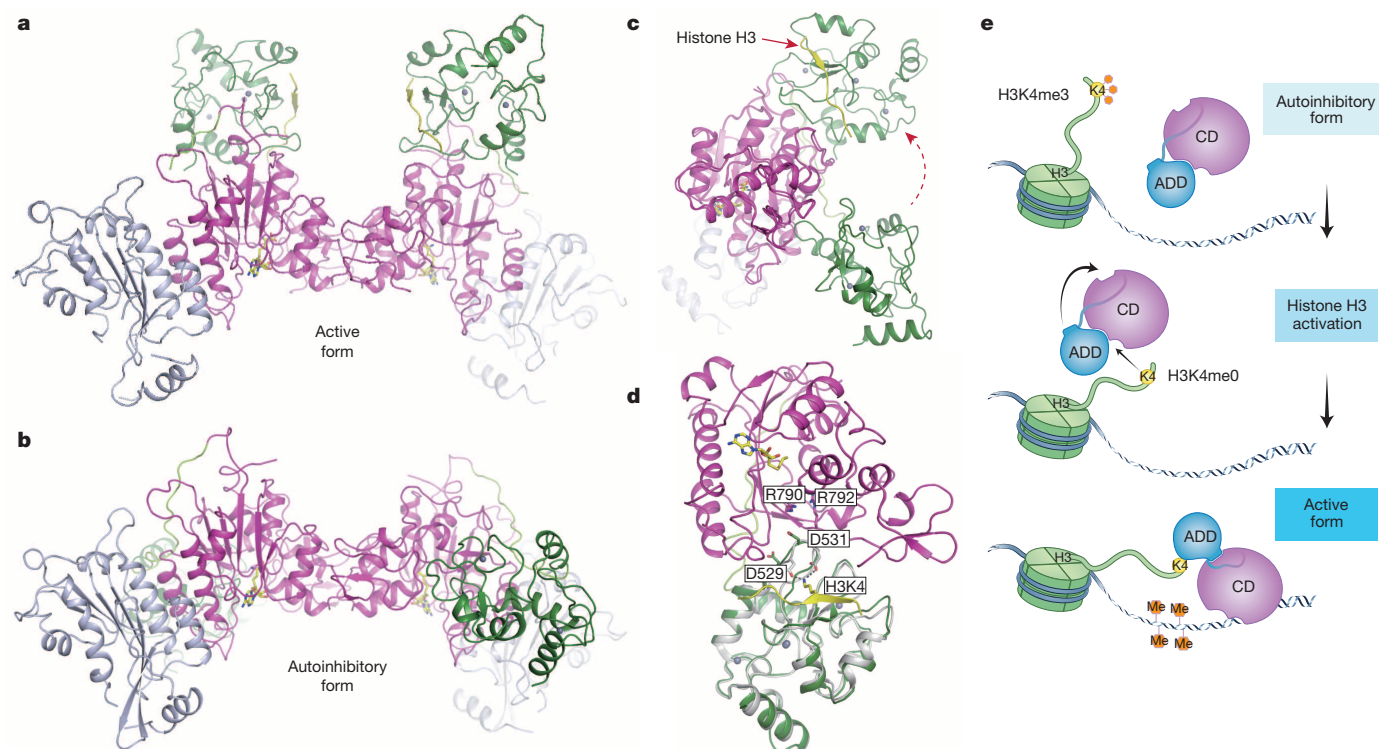


Figure 4 | Mechanisms for autoinhibition and histone H3 tail-induced activation of DNMT3A. **a**, Ribbon representations of the overall structure of the ADD-CD- C^{DNMT3L} -H3 complex. Histone H3 peptides are coloured in yellow. **b**, ADD-CD- C^{DNMT3L} structure in autoinhibitory form are shown for comparison. **c**, Ribbon representations of ADD-CD- C^{DNMT3L} (autoinhibitory form) and ADD-CD- C^{DNMT3L} -H3 (active form) complex structures with the CD domains superimposed. **d**, Superimposition of

ADD-CD- C^{DNMT3L} (C^{DNMT3L} omitted for simplicity) and H3-ADD structures. The ADD domain and histone H3 in the H3-ADD structure are coloured in grey and yellow, respectively. Residues for H3-ADD and ADD-CD interactions are shown in stick representation. **e**, A working model for the autoinhibition and histone H3 tail-induced activation of DNMT3A. DNMT3L and the PWWP domain of DNMT3A are not shown for simplicity.

induced by histone H3 (Fig. 4a–c and Extended Data Fig. 8a). DNMT3A may prefer an autoinhibitory form in the absence of histone H3. Upon histone H3 association, DNMT3A adopts an active form to open an otherwise locked DNA-binding region for establishment of DNA methylation. The negative-stain electron microscopy density maps clearly showed a distinct shape for DNMT3A2 (residues 275–912)- C^{DNMT3L} complex in the absence or presence of histone H3 peptide, supporting a conformational change of DNMT3A2- C^{DNMT3L} induced by histone H3: that is, transforming from an 'X' shape (autoinhibitory form) to a butterfly shape (active form) (Extended Data Fig. 8b–d). One-dimensional ^{19}F NMR measurement²⁵ also indicated a significant change of chemical shift (representing conformational change) for residue F827 (within the loop L2) upon histone H3 peptide (but not H3K4me3) titration, whereas only a slight change was observed for residue F868 (as a negative control) (Extended Data Fig. 8e, f). Collectively, both electron microscopy and NMR measurements support the existence of conformational change of DNMT3A induced by histone H3 tail.

Superimposition of H3-ADD (PDB accession number 3A1B) and ADD-CD- C^{DNMT3L} (autoinhibitory form) structures indicated that ADD domains in the two structures adopt a similar fold, and that histone H3 peptide in the H3-ADD structure has no overlap with the ADD-CD interface (Fig. 4d). Our previous studies have shown that mutation D529A or D531A not only significantly decreased H3-ADD interaction (Extended Data Fig. 7g) but also markedly decreased ADD-CD interaction (Fig. 2a), and released the autoinhibition (Fig. 2g, h). Moreover, H3K4me0 peptide could disrupt ADD-CD interaction (Fig. 2b). Thus, histone H3 releases the autoinhibition of DNMT3A through binding to residues of the ADD domain on the ADD-CD interface (such as D529 and D531) and disrupting the intramolecular interaction. The two residues function as a critical switch that can exist in both forms of DNMT3A and couple histone H3 recognition to the release of the autoinhibition.

Here we propose a working model for histone H3-induced dynamic regulation of the *de novo* DNA methyltransferase (Fig. 4e and Supplementary Video 1). DNMT3A exists in dynamic equilibrium between autoinhibitory and active forms, and the ADD domain oscillates between the two conformations. In the absence of histone H3, DNMT3A prefers an autoinhibitory form, in which the ADD domain binds to the CD domain and hinders its DNA-binding affinity. Once DNMT3A (or the DNMT3A-DNMT3L complex) is recruited to the nucleosome, H3K4me0 binds to the ADD domain and stimulates DNMT3A to undergo a significant conformational change from an autoinhibitory form to an active form. The ADD-CD interaction in the active form allows DNMT3A to adopt a relative stable conformation so that DNA methylation occurs in a range permitted by the histone H3 tail. Even when DNMT3A is recruited to an H3K4me3-containing nucleosome, the enzyme will remain in its autoinhibitory form to avoid DNA methylation within an impermeable chromatin environment. Our work reinforces the connection between DNA methylation and histone modifications, and sheds new light on the fine tuning of the establishment of DNA methylation. The structures may also provide a basis for the design of specific regulators for potential therapeutic applications.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 May; accepted 29 September 2014.

Published online 10 November 2014.

1. Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genet.* **33** (suppl.), 245–254 (2003).
2. Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nature Rev. Genet.* **14**, 204–220 (2013).

3. Law, J. A. & Jacobsen, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Rev. Genet.* **11**, 204–220 (2010).
4. Guibert, S., Forne, T. & Weber, M. Dynamic regulation of DNA methylation during mammalian development. *Epigenomics* **1**, 81–98 (2009).
5. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
6. Okano, M., Xie, S. & Li, E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genet.* **19**, 219–220 (1998).
7. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
8. Bestor, T. H. The DNA methyltransferases of mammals. *Hum. Mol. Genet.* **9**, 2395–2402 (2000).
9. Laurent, L. *et al.* Dynamic changes in the human methylome during differentiation. *Genome Res.* **20**, 320–331 (2010).
10. Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
11. Bird, A. P. & Taggart, M. H. Variable patterns of total DNA and rDNA methylation in animals. *Nucleic Acids Res.* **8**, 1485–1497 (1980).
12. Ehrlich, M. *et al.* Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res.* **10**, 2709–2721 (1982).
13. Chedin, F., Lieber, M. R. & Hsieh, C. L. The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc. Natl Acad. Sci. USA* **99**, 16916–16921 (2002).
14. Cedar, H. & Bergman, Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nature Rev. Genet.* **10**, 295–304 (2009).
15. Ooi, S. K. *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**, 714–717 (2007).
16. Otani, J. *et al.* Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX–DNMT3–DNMT3L domain. *EMBO Rep.* **10**, 1235–1241 (2009).
17. Zhang, Y. *et al.* Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res.* **38**, 4246–4253 (2010).
18. Li, B. Z. *et al.* Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase. *Cell Res.* **21**, 1172–1181 (2011).
19. Weber, M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genet.* **39**, 457–466 (2007).
20. Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770 (2008).
21. Chen, T., Ueda, Y., Xie, S. & Li, E. A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. *J. Biol. Chem.* **277**, 38746–38754 (2002).
22. Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A. & Cheng, X. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* **449**, 248–251 (2007).
23. Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. HhaI methyltransferase flips its target base out of the DNA helix. *Cell* **76**, 357–369 (1994).
24. Rajavelu, A., Jurkowska, R. Z., Fritz, J. & Jeltsch, A. Function and disruption of DNA methyltransferase 3a cooperative DNA binding and nucleoprotein filament formation. *Nucleic Acids Res.* **40**, 569–580 (2012).
25. Shi, P. *et al.* Site-specific protein backbone and side-chain NMR chemical shift and relaxation analysis of human vinexin SH3 domain using a genetically encoded ¹⁵N/¹⁹F-labeled unnatural amino acid. *Biochem. Biophys. Res. Commun.* **402**, 461–466 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank staff of beamline BL17U at Shanghai Synchrotron Radiation Facility, China, for their assistance in data collection, and H. Wang for help on electron microscopy analyses. We thank staff of the Biomedical Core Facility, Fudan University, for their help on biochemical analyses, and A. D. Riggs for providing the complementary DNAs of DNMT3A and DNMT3L. This work was supported by grants from the National Basic Research Program of China (2011CB965300, 2009CB918600, 2013CB910401), the National Science & Technology Major Project 'Key New Drug Creation and Manufacturing Program' of China (2014ZX09507-002, 2011ZX09506-001), the National Natural Science Foundation of China (31270779, 91419301, 31030019, U1432242, 31270771, 31222016, 31300685, U1332138), the Basic Research Project of Shanghai Science and Technology Commission (12JC1402700, 13JC1406300), the Fok Ying Tung Education Foundation (20090071220012), and the Chinese Academy of Sciences Pilot Strategic Science and Technology Projects B (numbers XDB08030201, XDB08030302). Y.C. is a scholar of the Hundred Talents Program of the Chinese Academy of Sciences.

Author Contributions X.G., L.W., and Y.X. designed the experiments. X.G., L.W., and X.Y. performed protein purification and crystallization of ADD–CD–C^{DNMT3L}–H3 complex; X.G. collected the data and determined the crystal structure. L.W., J.X., and S.H. performed protein purification, crystallization, and data collection of ADD–CD–C^{DNMT3L}. J.L. and J.W. determined the crystal structure. X.G. and L.W. performed enzymatic assays, fluorescence polarization and pull-down assays. L.D. and G.L. prepared nucleosomes for assays. C.T. and P.S. performed and analysed the ¹⁹F NMR measurements. Z.D. and Y.C. performed and analysed the electron microscopy measurements. X.G., L.W., and Y.X. analysed the data and wrote the manuscript. Y.X. supervised the project.

Author Information The coordinates and structure factors for the ADD–CD–C^{DNMT3L} and ADD–CD–C^{DNMT3L}–H3 structures have been deposited in the Protein Data Bank under accession numbers 4U7P and 4U7T, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.X. (xuyh@fudan.edu.cn).

METHODS

Protein expression and purification. The full-length DNMT3A2 (residues 224–912 of DNMT3A) was expressed in sf9 cells using the Bac-to-Bac system (Invitrogen). The infected cells were harvested and lysed in 50 mM Tris-Cl pH 8.0, 500 mM NaCl, and 0.01% 2-mercaptoethanol. The clarified lysate was applied onto GST affinity columns (GE Healthcare) and the fusion protein was cleaved with PreScission protease. The protein was stored in 20 mM Tris-Cl pH 8.0, 300 mM NaCl, and 0.01% 2-mercaptoethanol for the assays.

Truncations of DNMT3A2 were cloned into modified pGEX-6p-1 vector and the C^{DNMT3L} (residues 178–379) was inserted into modified pRSFDuet-1 vector. DNMT3A proteins were expressed independently or with C^{DNMT3L} in *Escherichia coli* strain BL21(DE3). The transformants were grown at 37 °C to an attenuation ($D_{600\text{ nm}}$) of 0.6 in 2× YT medium. The cultures were induced by adding 1 mM isopropyl-β-D-thiogalactopyranoside and further incubated for 16 h at 15 °C. The supernatant of cell lysate was applied onto GST affinity columns (GE Healthcare) and the fusion protein was digested with PreScission protease. The eluted protein was purified by ion exchange and gel filtration chromatography. The purified proteins were subjected to SDS–PAGE and stained by Coomassie blue. The peak fractions were concentrated to 5–10 mg ml^{−1} and used for crystallization and biochemical assays. Mutations of DNMT3A were purified in a similar procedure.

Crystallization and structure determination. For crystallization of the autoinhibitory form of DNMT3A, the complex of ADD–CD of DNMT3A (residues 455–912) and C^{DNMT3L} was mixed with a palindromic 18-base-pair DNA duplex (5′-G AGGCTAGCGCTAGCCTC-3′) and AdoHcy in a 1:1.2:2 molar ratio. The crystals were obtained using the hanging-drop, vapour-diffusion method by mixing 1 μl ADD–CD–C^{DNMT3L} complex with 1 μl reservoir solution containing 0.05 M Bis-Tris pH 5.6–6.0, 0.1 M sodium malonate and 8% PEG3350 at 4 °C. Crystals were cryoprotected by the reservoir buffer containing 22% glycerol and then flash frozen in liquid nitrogen. Although the DNA duplex was added for the crystallization, no corresponding electron density was observed. The DNA duplex may function as an additive to favour the crystallization.

For the crystallization of the active form of DNMT3A, the complex of ADD–CD of DNMT3A (residues 476–912) and C^{DNMT3L} was pre-incubated with histone H3 peptide (residues 1–12, ARTKQTARKSTG) at a 1:10 molar ratio before crystallization. Crystals were grown by the hanging-drop, vapour-diffusion method by mixing 1 μl protein (10 mg ml^{−1}) with 1 μl reservoir solution containing 100 mM sodium acetate (pH 5.3–5.6) and 600 mM ammonium sulphate at 18 °C. Crystals were cryoprotected by the reservoir buffer with 25% ethylene glycol and then flash frozen in liquid nitrogen.

The data were collected on beamline BL17U at Shanghai Synchrotron Radiation Facility, China, at wavelengths of 1.2816 Å and 0.9792 Å, respectively. Data were indexed, integrated, and scaled using the program HKL2000 (ref. 26). The orientation and position of CD–C^{DNMT3L} in the ADD–CD–C^{DNMT3L} complex was first determined by molecular replacement using CD–C^{DNMT3L} (PDB accession number 2QRV)²² as a searching model in the PHASER program²⁷. The resulting CD–C^{DNMT3L} model was refined with PHENIX package²⁸. The ADD domain then was put into the refined extra-difference density using MOLREP with the ‘search for model in the map’ module²⁹. The overall structure of the ADD–CD–C^{DNMT3L} complex was finally refined with stereochemistry and the reference structure ADD domain (PDB accession number 3A1B)¹⁶ as restraints. The anomalous Fourier map of zinc cations in the ADD domain confirmed the validity of the position of the ADD domain. The structure of the ADD–CD–C^{DNMT3L}–H3 complex was determined by molecular replacement using the ADD domain (PDB accession number 3A1B)¹⁶ and CD–C^{DNMT3L} (PDB accession number 2QRV)²² as searching models in the PHASER program²⁷, and was then manually built by COOT³⁰.

All refinements used the module phenix.refine of PHENIX²⁸. The model quality was checked with the PROCHECK program³¹. In the structure of ADD–CD–C^{DNMT3L}, 85.7% of residues were in most favoured regions, 12.9% in additional allowed regions, 0.9% in generously allowed regions, and 0.5% in disallowed regions. In the structure of ADD–CD–C^{DNMT3L}–H3, 88.4% of residues were in most favoured regions, 11% in additional allowed regions, and 0.6% in generously allowed regions. All structure figures were generated by PyMol³².

In vitro DNA methylation assay on naked DNA. The enzymatic activity of DNMT3A proteins was assessed by incorporation of a ³H-labelled methyl group from S-adenosyl-L-[methyl-³H]methionine ([methyl-³H]AdoMet, PerkinElmer)^{22,33}. A biotin-labelled DNA fragment amplified from the EBNA1 region of p220.2 (1.2 kilobases, 52 CG sites) was used as a substrate. For histone H3 stimulation, 0.3 μM DNMT3A proteins were pre-incubated with or without 3 μM histone H3 peptides (residues 1–12). For ADD-mediated inhibition, CD proteins (1 μM) were supplemented with or without the ADD domain proteins (18 μM) on ice for 30 min. DNA (100 ng) was methylated by DNMT3A proteins in the presence of 2.5 μM [methyl-³H]AdoMet, 25 mM Tris–HCl (pH 7.5), 5% glycerol, 0.01% 2-mercaptoethanol, and 0.5 mg ml^{−1} BSA. The reactions were incubated at 37 °C for 30 min, and terminated

by adding cold wash buffer (500 mM NaCl and 1 mM EDTA in PBST). The DNA products were immobilized on streptavidin beads, washed five times, and subjected to liquid-scintillation counting (PerkinElmer). Each reaction was performed in triplicate.

In vitro DNA methylation assay on recombinant poly-nucleosomes. Recombinant *Xenopus laevis* histones were expressed and purified as described³⁴. Site-specific methylation of H3K₄me3 was performed by the methyl-lysine analogs approach³⁵. Incorporation of H3K₄me3 was verified by specific antibodies (anti-H3K₄me3: Cell Signaling, 9751 s; anti-H3: Abcam, ab1791) and visualized on Tanon-5200 Chemiluminescent Imaging System (Tanon Science & Technology). Assembly of histone octamers and reconstitution of poly-nucleosomes were performed by salt dialysis using the 601 sequence (30 CG sites)³⁴. The reaction mixture was the same as described above except that 300 ng nucleosomal DNA was methylated using 1 μM DNMT3A proteins. The reaction was quenched by the addition of excess TE buffer and 1% SDS. The histone proteins were removed by phenol–chloroform extraction. Then DNA was purified by ethanol precipitation, resolved in TE, and subjected to liquid-scintillation counting (PerkinElmer). Each reaction was performed in triplicate.

Electrophoretic mobility-shift assay. A 6-carboxy-fluorescein (FAM)-labelled double-stranded DNA (dsDNA) containing one CpG site was generated from annealing two primers (upper primer, FAM-5′-CTGAATACTACTTGCGCTCTCTAACCTGAT-3′; lower primer, 5′-GACTTATGATGAACGCGAGAGATTG GACTA-3′). The FAM-labelled DNA was used both in electrophoretic mobility-shift assays and fluorescence polarization assays. DNA (25 nM) and the indicated amounts of proteins were incubated in reaction buffer containing 20 mM HEPES pH 7.5, 100 mM KCl, 8% glycerol, and 0.5 mg ml^{−1} BSA for 30 min at 25 °C. The samples were subject to a 12% PAGE and analysed by Typhoon FLA 9500 (GE Healthcare) image scanning.

Fluorescence polarization assay. FAM-labelled dsDNA (15 nM) was incubated with increasing amounts of DNMT3A proteins for 30 min at 25 °C in reaction buffer containing 20 mM HEPES pH 7.5, 100 mM KCl, 8% glycerol, and 0.5 mg ml^{−1} BSA. Fluorescence polarization measurements were performed on a Synergy 4 Microplate Reader (BioTek) at 25 °C. The bound fractions were calculated as (mP – baseline mP)/(maximum mP – baseline mP), in which mP (milli-polarization units) represents the fluorescence polarization value. For the peptide stimulation experiment, 15 nM DNA was pre-incubated with 0.2 μM CD–C^{DNMT3L} or 1 μM ADD–CD–C^{DNMT3L} protein. An increasing amount of histone H3 peptide (H3K₄me0 or H3K₄me3) was then added into the protein–DNA complex. The levels of protein–DNA complex formation were measured by fluorescence polarization. Each reaction was performed in triplicate. The curves were fitted using GraphPad Prism 5.

Isothermal titration calorimetry. To obtain the binding affinity between ADD–CD–C^{DNMT3L} and H3 peptide, 0.04 mM ADD–CD–C^{DNMT3L} (in cell) in the absence or presence of 0.5 mM dsDNA (30 base pairs, upper strand: 5′-CTGAATACTACTTGCGCTCTCTAACCTGAT-3′) was titrated with 0.5 mM histone H3 peptide (residues 1–12) (in syringe) using an iTC200 microcalorimeter (GE Healthcare) at 18 °C. Protein, DNA, and peptide were prepared in a buffer containing 10 mM HEPES, pH 8.0, 100 mM NaCl, and 0.5 mM TCEP. To obtain the binding affinity between the ADD domain and H3 peptide, 0.05 mM ADD protein (in cell) was titrated with 0.5 mM histone H3 peptide (residues 1–12) (in syringe). The data were fitted by Origin 7.0 software.

In vitro binding assay. For the GST pull-down assay, 30 μg DNMT3A–CD proteins were incubated with 8 μg GST–ADD–linker fusion proteins for 1 h at 4 °C in binding buffer containing 20 mM Tris–HCl pH 8.0, 100 mM NaCl, 0.01% 2-mercaptoethanol, 5% glycerol, and 0.1% Triton X-100. GST–ADD–linker proteins were then immobilized to 25 μl of glutathione resins (GE Healthcare) for 1 h at 4 °C. After washing three times with binding buffer, bound proteins were subjected to SDS–PAGE and stained by Coomassie blue.

For the histone peptide pull-down assay, 1 μg biotinylated histone H3 peptide (residues 1–21) was incubated with 30 μg wild-type and mutant ADD–CD–C^{DNMT3L} proteins for 1 h at 4 °C in binding buffer containing 20 mM Tris–HCl pH 8.0, 250 mM NaCl, 0.01% 2-mercaptoethanol, 5% glycerol, and 0.1% Triton X-100. Then 20 μl streptavidin beads were added into the mixture and incubated 1 h at 4 °C. After washing three times with binding buffer, bound proteins were subjected to SDS–PAGE and stained by Coomassie blue.

Expression of ¹⁹F-labelled proteins. An orthogonal tRNA/tRNA synthetase system was used to incorporate ¹⁹F-labelled unnatural amino acid. Briefly, a TAG stop codon was introduced into the desired site to encode L-4-trifluoromethylphenylalanine (tfmF). A modified pEVOL-tfmFRS plasmid to express tRNA_{CUA} and tfmF-specific aminoacyl-tRNA was co-transformed with the TAG-carrying plasmid into BL21 (DE3)³⁶. The bacterial culture was induced with 0.02% L-arabinose, 1 mM tfmF, and 1 mM IPTG. The ¹⁹F-labelled proteins were purified as for wild-type DNMT3A proteins.

¹⁹F NMR spectra measurements. All one-dimensional ¹⁹F NMR spectra measurements were performed at 293 K on an Agilent 500 MHz spectrometer equipped

with an HFT probe, and the observation channel was tuned to ^{19}F (470.2 MHz), with 1,024 free induction decay accumulations in every 4 s recycling delay. A one-dimensional ^{19}F spectrum was acquired with one pulse program with a 90° pulse width of 12.45 μs and power at 57 W. The spectrum width was 60 p.p.m. and offset was -62 p.p.m. ^{19}F chemical shifts were referenced to an external standard, tfmF (-62.38 p.p.m.). The data were processed and plotted with an exponential window function (line broadening = 20 Hz) using ACD/NMR Processor Academic Edition software (ACD/Labs). The spectra of 0.14 mM ADD-CD F686tfmF or 0.12 mM ADD-CD F827tfmF with or without histone H3 peptides (residues 1–12) were collected at 293 K.

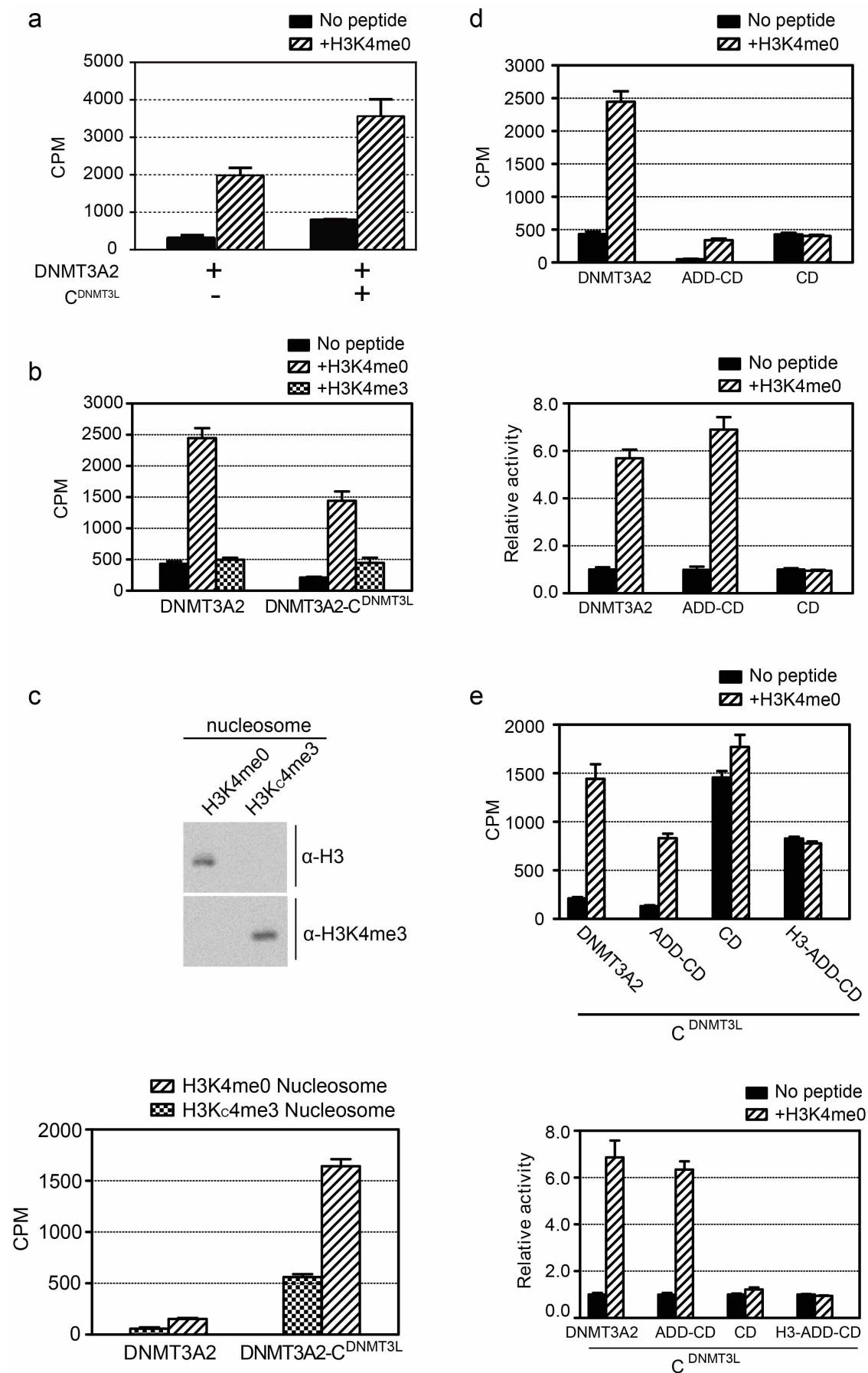
Electron microscopy data collection. DNMT3A2- C^{DNMT3L} in the absence or presence of histone H3 peptide was analysed by negative-stain electron microscopy in the same manner. The samples were prepared by dilution of purified protein complex to $6.15\ \mu\text{g ml}^{-1}$, then 4–5 μl of this sample was deposited onto a glow-discharged 400 mesh continuous carbon grid (Beijing Zhongjingkeyi Technology). The sample was then stained with 2% uranyl formate and air-dried. Data were recorded on a Tecnai G2 F20 TWIN transmission electron microscope (FEI) equipped with a field-emission gun operated at 200 kV. Images were recorded at $\times 71,000$ microscope magnification on a $4\text{k} \times 4\text{k}$ Eagle CCD (charge-coupled device) camera with a pixel size of 1.15 Å per pixel. The defocus ranged from -0.5 to $-0.8\ \mu\text{m}$.

Electron microscopy image processing and three-dimensional reconstruction. For this, 9,864 and 21,910 particles were boxed out for DNMT3A2- C^{DNMT3L} and DNMT3A2- C^{DNMT3L} -H3, respectively, by using the e2boxer.py program in EMAN2.1 (ref. 37). Contrast transfer function parameters were determined for particles boxed out from each CCD image using EMAN1.9 procedure ctfit, followed by phase flipping using the applyctf program. The data were then low-pass filtered to 10 Å to enhance the image contrast for three-dimensional reconstruction^{38,39}. Reference-free two-dimensional analysis used the EMAN1.9 program refine2d.py and IMAGIC⁴⁰, and those class-averages were used to generate initial models by e2initialmodel.py in EMAN2.1. Three-dimensional reconstruction was performed by the EMAN1.9 program refine^{41,42}. Initially, no symmetry was imposed in the reconstruction process, and the resulting three-dimensional reconstruction revealed the existence of a two-fold symmetry in both maps but in different locations, which was subsequently imposed in the reconstruction process. The final resolution was estimated at 24 Å and 20 Å, respectively, by the 0.5 FSC criteria using the eotest program in EMAN1.9.

UCSF Chimera (<http://www.cgl.ucsf.edu/chimera/>)⁴³ was used to render the electron microscopy density together with the crystal structures. In addition, the Fit In

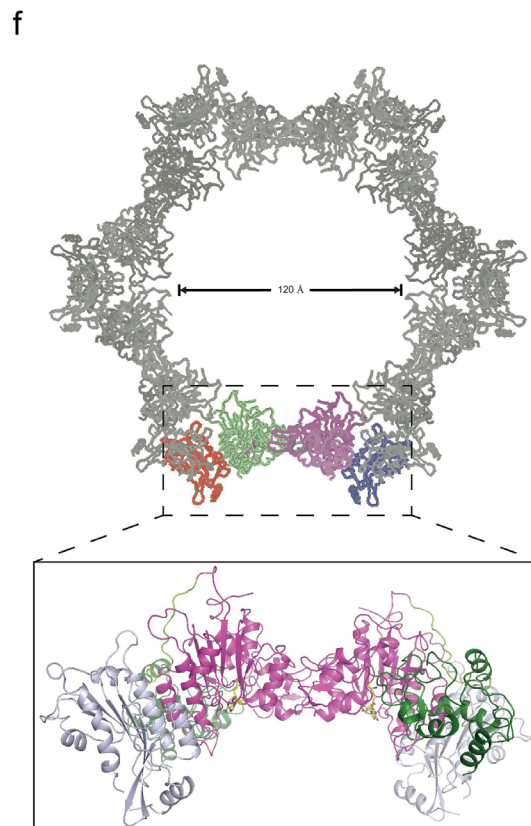
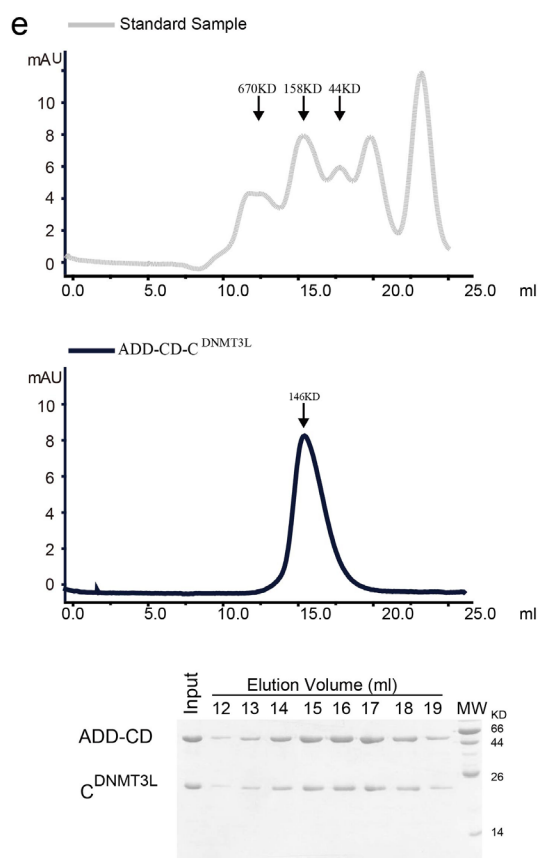
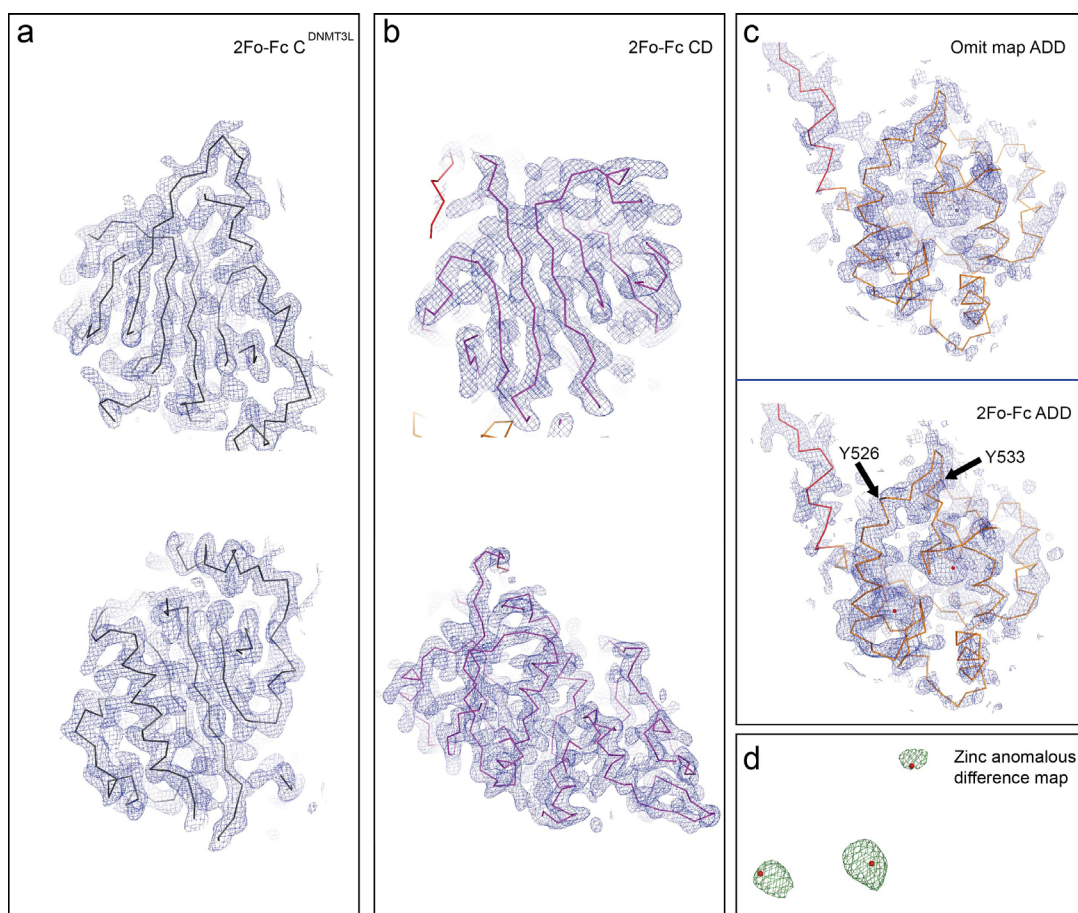
Map module in Chimera was used for rigid-body fitting of the crystal structure into the corresponding electron microscopy density map.

26. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
27. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
28. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).
29. Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022–1025 (1997).
30. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
31. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
32. DeLano, W. L. The PyMOL Molecular Graphics System. <http://www.pymol.org> (2002).
33. Roth, M. & Jeltsch, A. Biotin-avidin microplate assay for the quantitative analysis of enzymatic methylation of DNA by DNA methyltransferases. *Biol. Chem.* **381**, 269–272 (2000).
34. Luger, K., Rechsteiner, T. J. & Richmond, T. J. Expression and purification of recombinant histones and nucleosome reconstitution. *Methods Mol. Biol.* **119**, 1–16 (1999).
35. Simon, M. D. *et al.* The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* **128**, 1003–1012 (2007).
36. Shi, P. *et al.* Site-specific (1)(9)F NMR chemical shift and side chain relaxation analysis of a membrane protein labeled with an unnatural amino acid. *Protein Sci.* **20**, 224–228 (2011).
37. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
38. Ludtke, S. J., Baldwin, P. R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82–97 (1999).
39. Ludtke, S. J., Jakana, J., Song, J. L., Chuang, D. T. & Chiu, W. A 11.5 Å single particle reconstruction of GroEL using EMAN. *JMB* (2001).
40. van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R. & Schatz, M. A new generation of the IMAGIC image processing system. *J. Struct. Biol.* **116**, 17–24 (1996).
41. Cong, Y. & Ludtke, S. J. Single particle analysis at high resolution. *Methods Enzymol.* **482**, 211–235 (2010).
42. Cong, Y. *et al.* 4.0-Å resolution cryo-EM structure of the mammalian chaperonin TRiC/CCT reveals its unique subunit arrangement. *Proc. Natl Acad. Sci. USA* **107**, 4967–4972 (2010).
43. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).



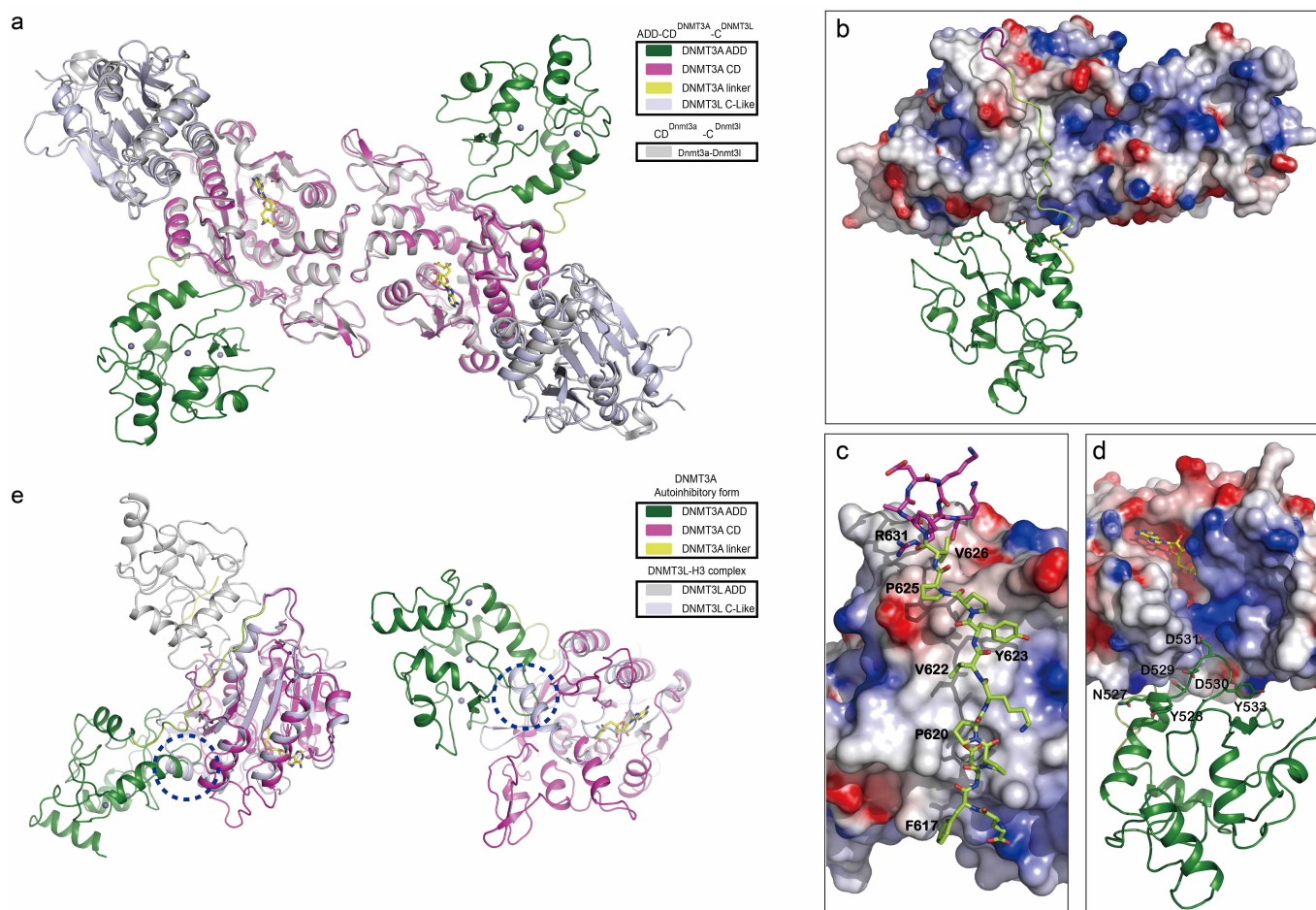
Extended Data Figure 1 | *In vitro* DNA methyltransferase activity of DNMT3A. **a**, *In vitro* DNA methyltransferase activities of DNMT3A2 (purified from insect cells) in the absence or presence of C^{DNMT3L}. The assays were performed in the presence or absence of histone H3 peptide (residues 1–12). Note that C^{DNMT3L} could enhance the activity of DNMT3A2 by a factor of 2–3, which is consistent with previous study²². However, histone H3-mediated activation of DNMT3A is independent of the existence of C^{DNMT3L}. **b**, *In vitro* DNA methyltransferase activities of DNMT3A2 (purified from insect cells) or DNMT3A2–C^{DNMT3L} (purified from bacteria) in the presence or absence of histone H3 peptides. **c**, Enzymatic activities of DNMT3A2 (purified from insect cells) or DNMT3A2–C^{DNMT3L} (purified from bacteria) using reconstituted nucleosomes as substrates. Nucleosomes containing unmodified

histone H3 or H3K₄me₃ were subject to SDS–PAGE and visualized using specific antibodies. **d**, **e**, Enzymatic activities of various N-terminal deletions of DNMT3A2 in the absence (**d**) or presence (**e**) of C^{DNMT3L}. Corresponding relative activities are indicated at the bottom of each figure. CPM, counts per minute. Error bars, s.d. for triplicate experiments. The ADD–CD or CD protein purified from bacteria was not stable in solution and tended to precipitate out, which may have resulted in their lower activities under our experimental conditions (compared with DNMT3A2 purified from insect cells). Because C^{DNMT3L} could stabilize DNMT3A and had no effect on histone H3-mediated activation, protein complexes ADD–CD–C^{DNMT3L} and CD–C^{DNMT3L} were used in the following studies if not specified.



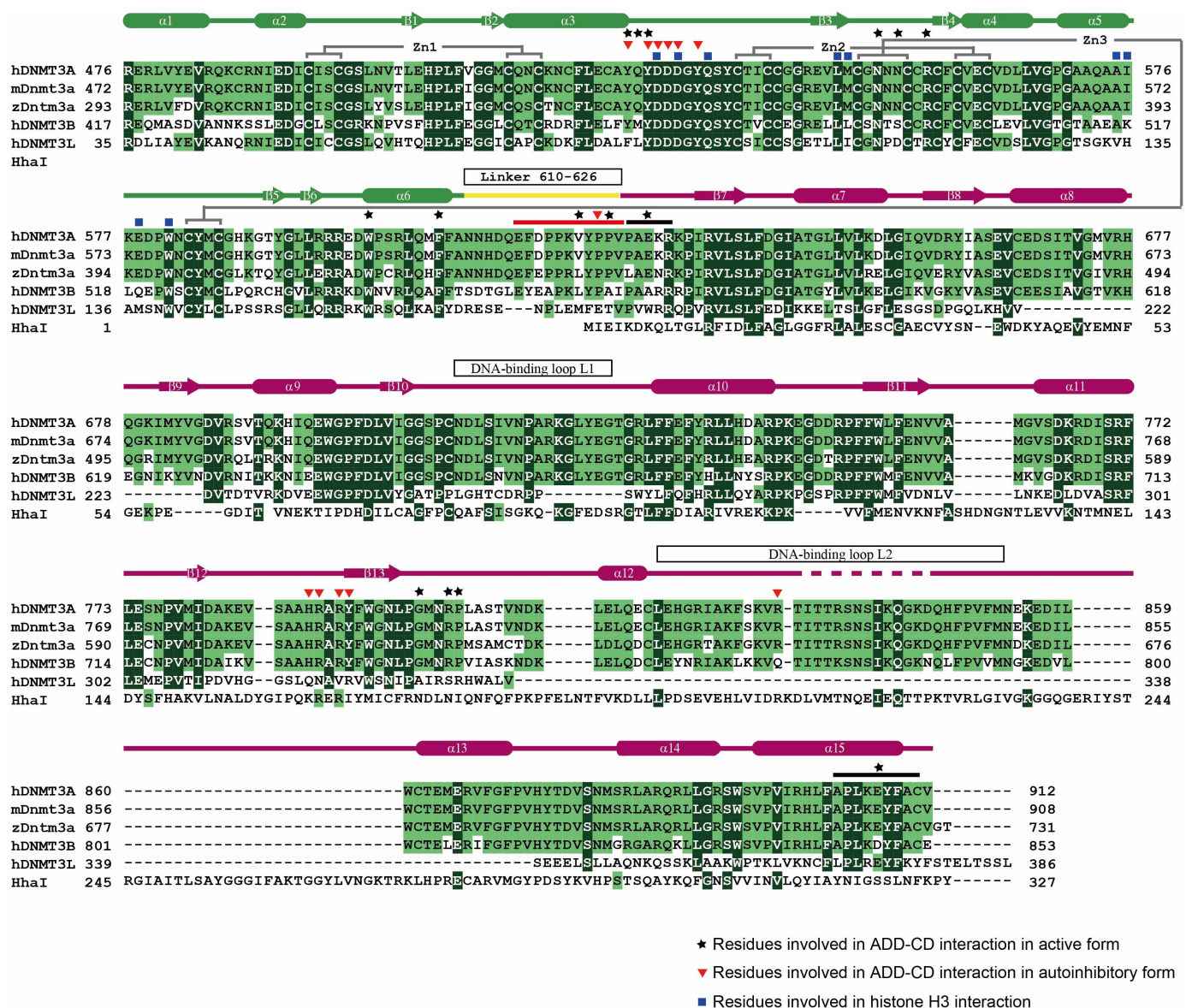
Extended Data Figure 2 | Crystal structure of ADD-CD- C^{DNMT3L} in autoinhibitory form. **a, b**, Two different views of the $2F_{\text{observed}} - F_{\text{calculated}}$ maps for C^{DNMT3L} (**a**) and CD (**b**) domains in the ADD-CD- C^{DNMT3L} structure. The maps were calculated at 3.82 Å and contoured at 1.5 σ . Only main-chains are shown for simplicity. **c**, The $2F_{\text{observed}} - F_{\text{calculated}}$ maps for the ADD domain after refinement of the CD- C^{DNMT3L} complex (top) and after refinement of the ADD-CD- C^{DNMT3L} complex (bottom). The maps were calculated at 3.82 Å and contoured at 0.8 σ . Main-chains from most residues, including residues 526–533 involved in the interaction with CD domain, fit well into the electron density. Some loop regions were not well covered by electron density, which is consistent with a high B factor (Extended Data Fig. 8a) of the ADD domain in the complex structure, supporting the dynamic feature of the ADD domain for regulating enzymatic activity of DNMT3A.

d, Zn-anomalous difference map contoured at 3.5 σ shows the positions of zinc cations in the ADD domain. **e**, Gel filtration profiles for standard proteins and the ADD-CD- C^{DNMT3L} complex. The peak position corresponds to the dimer of ADD-CD- C^{DNMT3L} with a molecular mass of about 140 kDa. **f**, Dimer formation of the ADD-CD- C^{DNMT3L} complex in crystals. The dimer of ADD-CD- C^{DNMT3L} complexes is mediated by CD-CD interaction in a two-fold crystallographic symmetry. Given the difficulty in tracing the conformation of the side chain in 3.82 Å resolution structure, we have not discussed the specific hydrogen bond or hydrophobic interaction within ADD-CD- C^{DNMT3L} . Residues 832–846 of DNMT3A were not built in the model because they lacked electron density, which may have resulted from their flexibility in crystals.



Extended Data Figure 3 | Structure of ADD-CD-C^{DNMT3L} in autoinhibitory form. **a**, Superimposition of human ADD-CD-C^{DNMT3L} with mouse CD^{DNMT3A}-C^{DNMT3L} (lack of ADD domain, PDB accession number 2QRV)²² structures shown in ribbon representations. CD-C^{DNMT3L} in two structures is well aligned with a root mean squared deviation of 1.28 Å for 723 Cα aligned. The function of DNMT3A-DNMT3L complex dimerization has been characterized in a previous study²². The functions and structures of the CD and C^{DNMT3L} domains, and the CD-CD and CD-C^{DNMT3L} interfaces, were not discussed in this work. **b**, Overall structure of ADD-CD-C^{DNMT3L} with CD-C^{DNMT3L} shown in electrostatic potential surface, and the ADD domain and linker shown in ribbon representation. The linker packs against a hydrophobic surface of the CD domain. **c**, **d**, Close-up view of linker-CD (c) and ADD-CD (d) interfaces with the electrostatic potential surface of the

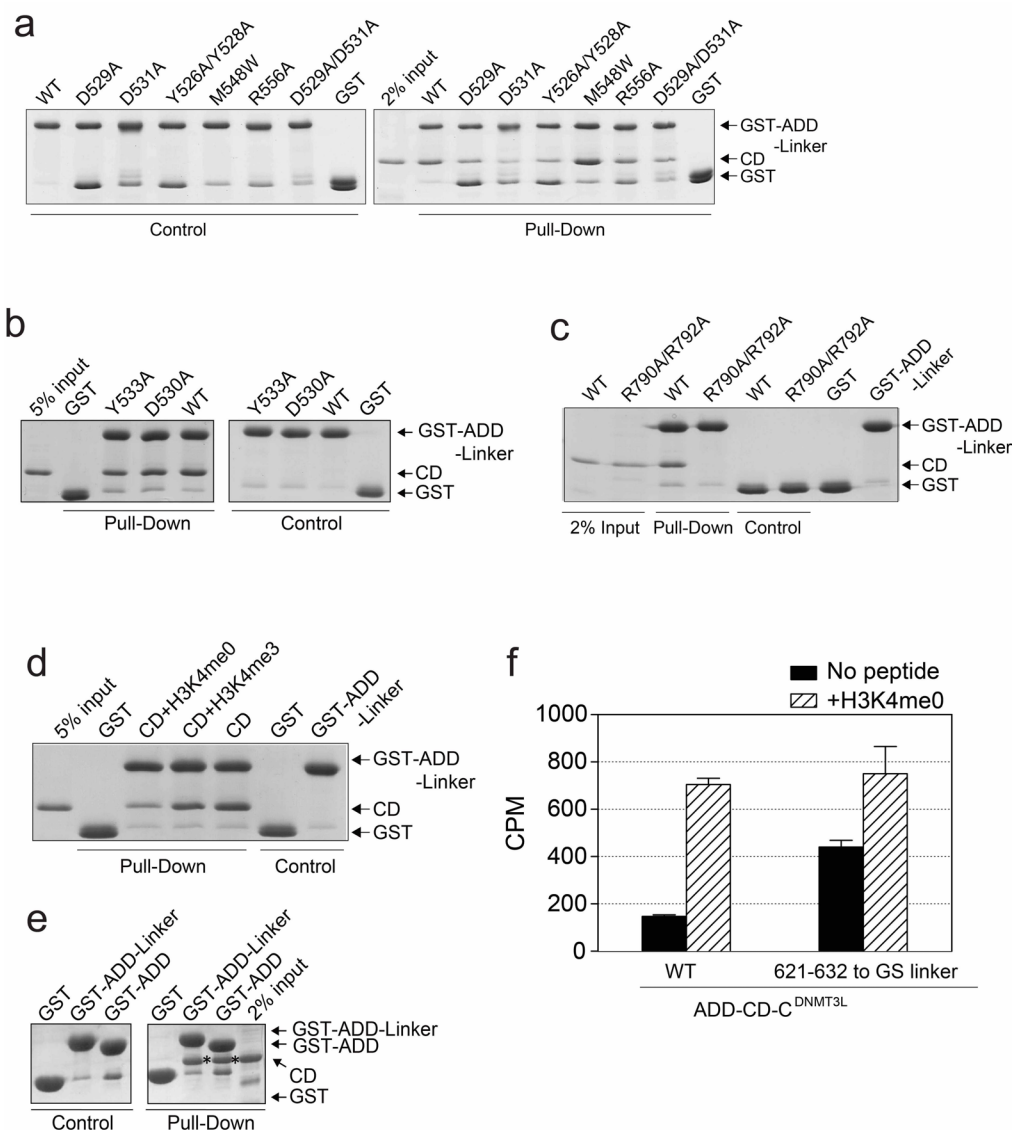
CD domain indicated. Critical residues are shown in stick representation. **e**, Superimposition of ADD-CD-C^{DNMT3L} and DNMT3L-H3 structures (PDB accession number 2PVC)¹⁵ shown in ribbon representations in two different views. The CD domain and C-like domain of DNMT3L were aligned for comparison. Note that the extended loop of the ADD domain in ADD-CD-C^{DNMT3L} overlaps with an α helix in the DNMT3L-H3 structure. DNMT3L is unlikely to adopt a similar conformation to that of ADD-CD-C^{DNMT3L} because otherwise the ADD domain will have steric hindrance with the C-like domain of DNMT3L (dashed circle). According to the above analyses, the structure of the autoinhibitory form of DNMT3A could not be predicted on the basis of the DNMT3L structure because the overall structures of DNMT3A and DNMT3L are different.



Extended Data Figure 4 | Sequence alignment of DNMT3 family members.

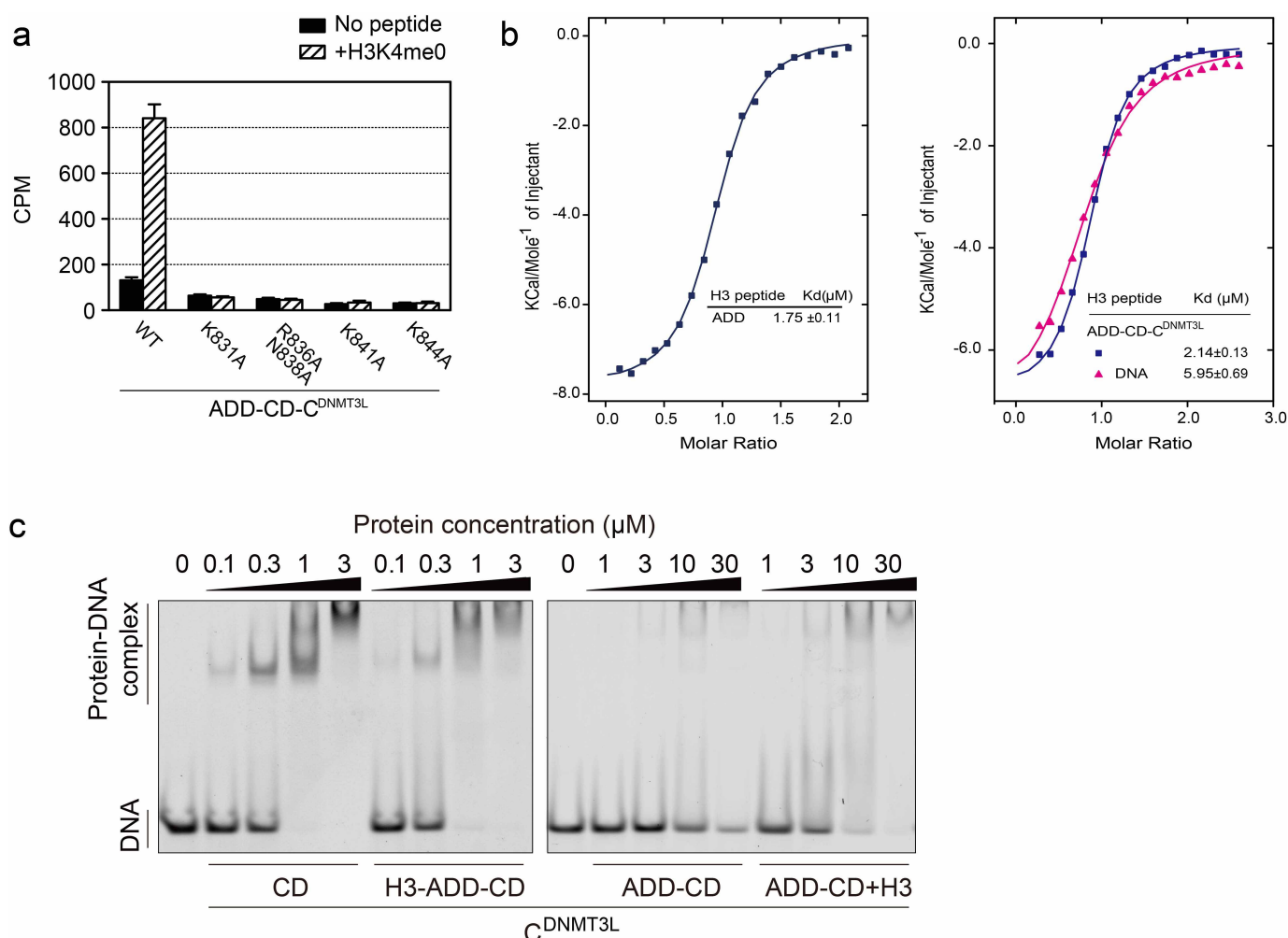
Sequences of human DNMT3A (NP_072046), DNMT3B (NP_008823), DNMT3L (NP_787063), mouse Dnmt3a (NP_031898), zebrafish Dnmt3a (NP_001018150), and DNA methyltransferase from *Haemophilus parahaemolyticus* (WP_005706946) used in the alignment. Highly conserved and identical residues are highlighted with dark green background, and conserved residues are indicated with light green background. Secondary

structural elements are coloured as in Fig. 1a and indicated above the sequences. Invisible residues in the structure of ADD-CD-^{DNMT3L} are indicated as dashed lines above the sequences. Residues involved in ADD-CD interactions in active form or autoinhibitory form are indicated as black stars and red triangles, respectively. Residues involved in H3-ADD interactions are indicated as blue squares.



Extended Data Figure 5 | Interactions between the ADD and CD domains. **a, b**, GST pull-down assays with recombinant CD (residues 627–912) protein incubated with wild-type or mutant GST-ADD-linker proteins immobilized on glutathione resin. The bound proteins were analysed by SDS-PAGE and Coomassie blue staining. **c**, GST pull-down assays using wild-type or mutant of the CD domain. **d**, GST pull-down assays in the absence

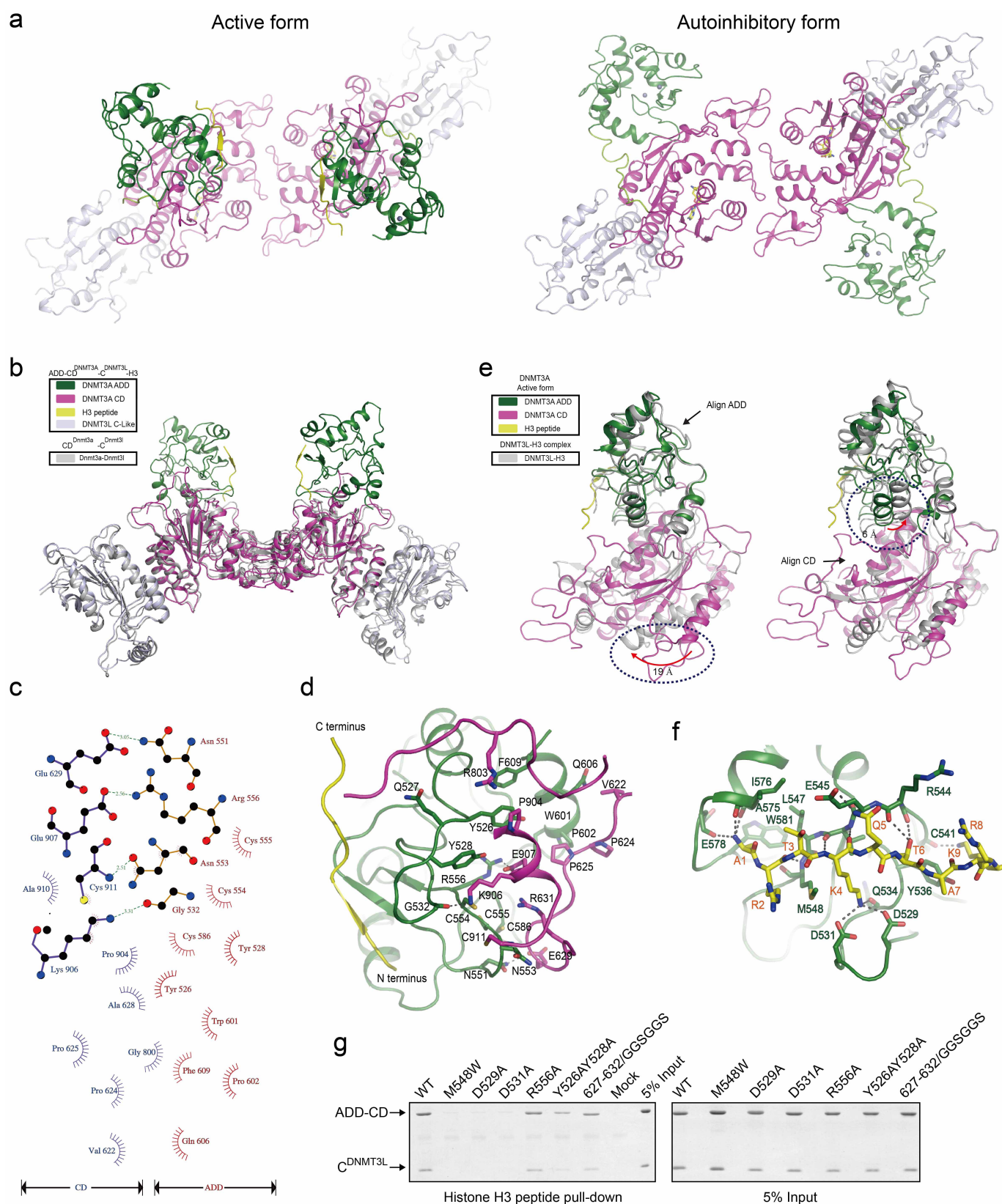
or presence of histone H3 peptide (H3K4me0 or H3K4me3). **e**, GST pull-down assays with the CD domain incubated with GST-ADD or GST-ADD-linker proteins immobilized on glutathione resin. **f**, Activities of wild-type and mutant ADD-CD. Residues 621–632 were replaced by a GS linker in the mutant proteins.



Extended Data Figure 6 | Interactions between DNMT3A and DNA.

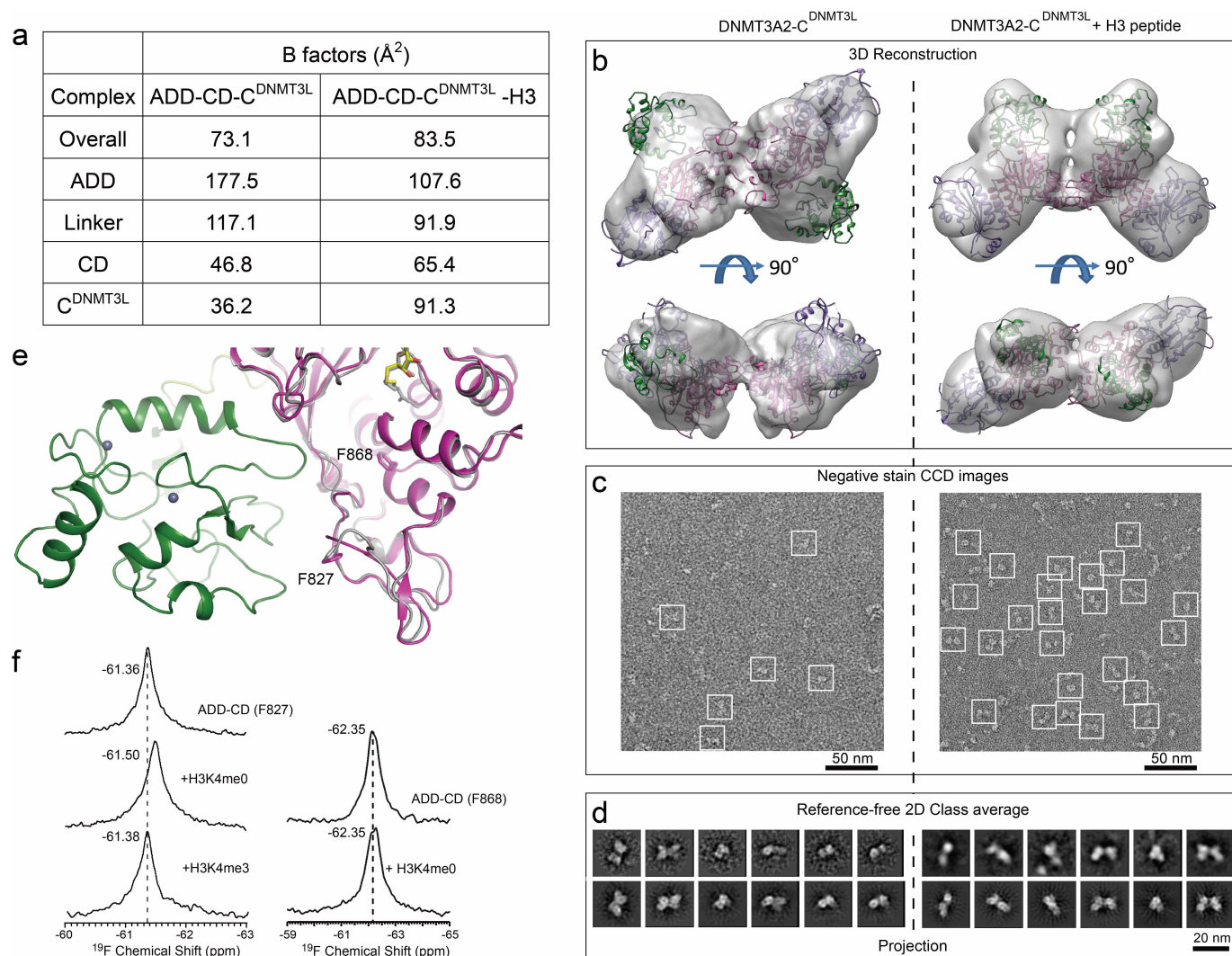
a, Enzymatic activities of wild-type and mutant ADD-CD-C^{DNMT3L}. Residues on the missing loop (residues 831–846) were mutated for the *in vitro* DNA methyltransferase activity assay. Error bars, s.d. for triplicate experiments. Mutating above residues leads to loss of activity of ADD-CD-C^{DNMT3L}, supporting their important role in catalysis or DNA recognition. The missing loop in the ADD-CD-C^{DNMT3L} structure is equivalent to a DNA-binding loop in the HhaI-DNA structure. **b**, DNA has no effect on the interaction between histone H3 and DNMT3A. Left, isothermal titration calorimetry enthalpy plot for the binding of isolated ADD domain (in cell) to histone H3 peptide (residues 1–12, in syringe), with the estimated binding affinities (K_d) listed. Right, superimposed isothermal titration calorimetry enthalpy plots for the binding of ADD-CD-C^{DNMT3L} (in cell) to histone H3 peptide (residues 1–12, in syringe) in the absence or presence of dsDNA. The estimated

binding affinities (K_d) are listed. Histone H3 peptide has comparable binding affinity to the ADD domain alone (1.75 μM) and ADD-CD-C^{DNMT3L} in autoinhibitory form (2.14 μM), and the addition of DNA was not able to enhance the binding affinity further. The presence of DNA led to a slight decrease in the binding affinity between histone H3 peptide and ADD-CD-C^{DNMT3L}, which may have resulted from slight precipitation of the protein caused by the high concentration of DNA used for titration. **c**, Electrophoretic mobility-shift assays for DNMT3A proteins in the absence or presence of histone H3 peptide, with protein concentrations indicated. H3-ADD-CD represents a fusion protein with histone H3 (residues 1–20) at the N terminus of ADD-CD. The assays showed that CD-C^{DNMT3L} strongly bound to the FAM-labelled DNA duplex, whereas the existence of the ADD domain markedly decreased DNA-binding affinity, which was partly restored by the addition of histone H3 peptide or largely restored by H3-ADD-CD fusion protein.



Extended Data Figure 7 | Structure of ADD-CD- C^{DNMT3L} -H3 in active form. **a**, Ribbon representations of the overall structures of the ADD-CD- C^{DNMT3L} in active (left) and autoinhibitory (right) forms. Histone H3 peptides are coloured in yellow. **b**, Structural comparison of human ADD-CD- C^{DNMT3L} -H3 and mouse CD^{Dnmt3a}- C^{Dnmt3L} complexes. The compared structures are shown in ribbon representations. The ADD-CD- C^{DNMT3L} complex structure (this study) is coloured as in Fig. 1d, and the CD^{Dnmt3a}- C^{Dnmt3L} complex structure²² is coloured in grey. Residues 611–620 and 833–846 of DNMT3A were not built in the model because they lacked electron density. **c**, LIGPLOT representation of the ADD-CD interactions in the ADD-CD- C^{DNMT3L} -H3 structure. Carbon, oxygen, and nitrogen are shown as black, red, and blue balls, respectively. Hydrogen bonds are indicated as dashed lines, with lengths given in Å. **d**, Close-up view of the ADD-CD interface. Critical residues for the interactions are shown in stick representation, and hydrogen bonds are indicated as dashed lines. The C terminus (residues 903–911) of the CD domain and a loop region (residues 621–632) together form a flat patch for interaction with the ADD domain. Hydrogen bonds are formed between residues N551, N553, and R556 of the

ADD domain and residues E629, C911, and E907 of the CD domain. Residues Y526, Y528, W601, and F609 of the ADD domain, V622 and P625 of the linker, and R803 and P904 of the CD domain are involved in hydrophobic interactions. **e**, Structural comparison of ADD-CD-H3 in ADD-CD- C^{DNMT3L} -H3 (this study) and DNMT3L-H3 structures (PDB accession number 2PVC)¹⁵. Two compared structures are shown in ribbon representations with ADD domains (left) or catalytic domains (right) aligned, respectively. The DNMT3L-H3 structure is coloured in grey. When the ADD domains are superimposed, the catalytic domain moves with a longest distance of 19 Å. When the CD domains are superimposed, the ADD domain moves 6 Å. **f**, Close-up view of the H3-ADD interface. Critical residues for the interactions are shown in stick representation, and hydrogen bonds are indicated as dashed lines. The fashion of histone H3-ADD interaction is similar to that observed in the structure of the H3-ADD fusion protein¹⁶. **g**, Histone H3 peptide pull-down assay. Recombinant wild-type and mutant ADD-CD- C^{DNMT3L} proteins were incubated with biotinylated histone H3 peptide (residues 1–21) and immobilized onto streptavidin sepharose beads. Bound proteins were subjected to SDS-PAGE and stained by Coomassie blue.



Extended Data Figure 8 | Conformational change of DNMT3A induced by histone H3 tail. **a**, Average B factors for domains of ADD-CD-C^{DNMT3L} in the structures of ADD-CD-C^{DNMT3L} and ADD-CD-C^{DNMT3L} bound to H3 peptide. The average B factor of the ADD domain is higher than other domains in both structures, and is higher in autoinhibitory form (177.5 \AA^2) than that in active form (107.6 \AA^2). The results indicate that the ADD domain is more dynamic than other domains of the complex, especially in its autoinhibitory form. The observation further supports the idea that DNMT3A undergoes conformational changes on the ADD domain induced by histone H3. **b**, Two different views of the electron microscopy density maps of DNMT3A2-C^{DNMT3L} (left) and DNMT3A2-C^{DNMT3L}-H3 (right) processed to 24 \AA and 20 \AA resolution, respectively. The corresponding crystal structure was fitted into the electron microscopy density map for each state. The density is not fully occupied, which might be because of the missing PWWP domain in the crystal structures. **c**, Typical negative stain CCD images of DNMT3A2-C^{DNMT3L} (left) and DNMT3A2-C^{DNMT3L}-H3 (right). Representative particles are highlighted

by white boxes. **d**, Comparison of the two-dimensional projections (bottom) from the electron microscopy map with the corresponding reference-free two-dimensional class averages (top) reveals similar structural features. **e**, Position of residues F827 and F868 for ^{19}F NMR measurements. Close-up view of the DNMT3A structure in autoinhibitory form with residues F827 and F868 indicated in stick representation. Residue F827 is located in loop L2 (for DNA binding) and close to the ADD domain. As a negative control, residue F868 is located close to the catalytic cavity and away from the ADD domain. Residue F868 is unlikely to undergo conformational change when the ADD domain dissociates from the CD domain. To detect conformational changes of DNMT3A in solution, residues F827 and F868 were substituted by ^{19}F -labelled L-4-trifluoromethylphenylalanine (^{19}F -tfmF) in ADD-CD. **f**, One-dimensional ^{19}F NMR measurements were performed using ADD-CD with substitution of F827tfmF (left) or F868tfmF (right) in the absence or presence of H3K4me0 or H3K4me3 peptide. The chemical shift for each measurement is indicated.

Extended Data Table 1 | Data collection and refinement

	ADD-CD-C ^{DNMT3L} complex	ADD-CD-C ^{DNMT3L} complex with H3 peptide
Data collection		
Space group	P 6 ₃ 2 2	P 6 ₅
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	252.0, 252.0, 75.3	183.8, 183.8, 123.3
α , β , γ (°)	90, 90, 120	90, 90, 120
Resolution (Å)	50.0 – 3.82 (3.96 – 3.82)*	50.0 – 2.90 (3.00 – 2.90)
<i>R</i> _{sym} or <i>R</i> _{merge}	0.098 (0.870)	0.095 (0.721)
<i>I</i> / σI	13.2 (1.9)	20.5 (3.5)
Completeness (%)	99.9 (99.9)	99.0 (100.0)
Redundancy	5.7 (5.8)	9.9 (9.8)
Refinement		
Resolution (Å)	50.0 – 3.82 (3.96 – 3.82)*	50.0 – 2.90 (3.00 – 2.90)
No. reflections	20564	52407
<i>R</i> _{work} / <i>R</i> _{free}	0.230 / 0.273	0.223 / 0.261
No. atoms		
Protein	5043	9820
H3 peptide		158
Ligand/ion	29	58
Water		11
B-factors (Å ²)		
Protein	73.2	83.2
H3 peptide		108.7
Ligand/ion	55.1	61.1
Water		60.5
R.m.s deviations		
Bond lengths (Å)	0.005	0.010
Bond angles (°)	0.92	1.40

* Highest resolution shell is shown in parentheses.

Extended Data Table 2 | Effect of DNMT3A mutants on autoinhibition

Mutations	Description	Effect
D529A	H3K4me0 recognition (residue K4) ADD-CD interface (autoinhibitory form)	Release
D531A	H3K4me0 recognition (residue K4) ADD-CD interface (autoinhibitory form)	Release
M548W	H3K4me0 recognition (residue A1)	Inhibition
Y526A/Y528A	ADD-CD interface (autoinhibitory form)	Release
Residues 627-632 replaced by GGSGGS	ADD-CD interface (in both forms)	Release
E907A	ADD-CD interface (active form)	No-change
R742G/P743G/K744S	Potentially for DNA recognition	No-change
K812A	Potentially for DNA recognition	No-change
R720G/K721S	Potentially for DNA recognition	Decrease
R836A/N838A	Potentially for DNA recognition	Decrease
K831A	Potentially for DNA recognition	Decrease
K841A	Potentially for DNA recognition	Decrease
K844A	Potentially for DNA recognition	Decrease
Residues 809-813 replaced by GGSGG	Potentially for DNA recognition	Decrease
Residues 821-846 replaced by GGSGGSGG	Potentially for DNA recognition	Decrease

All mutants were made on ADD-CD (residues 476–912) and the protein complexes ADD-CD-^{CDNMT3L} were used for the assays. 'Release' represents release of the autoinhibition by at least twofold activity enhancement. 'Inhibition' represents inhibition of activity and no response to H3 peptides. 'No-change' indicates that the mutants behaved similar to wild-type protein. 'Decrease' represents decrease or loss of enzymatic activity.

CAREERS

CAREER PATHWAYS Building a database of possibilities **p.647**

GRANTS Agency limits awards for the well-funded **p.647**

NATUREJOBS For the latest career listings and advice www.naturejobs.com



ANDREW BAKER/GETTY

CAREER COUNSELLING

Pick a path

Where to go to get advice on finding a job.

BY NEIL SAVAGE

Sarah Cullen was eyeing the impending end of her postdoctoral position and wondering what to do next. She had earned a PhD in microbiology and immunology at the University of Arkansas in Fayetteville in 2009, and was now studying the behaviour of breast-cancer cells. But she felt that it was time for a change: she knew that postdocs are jammed into the scientific pipeline every year but fill only a tiny number of faculty research jobs (see *Nature* **511**, 255–256; 2014).

“The career options were limited, and I saw so few friends getting academic positions,” Cullen says. “Rather than being a 20-year postdoc, I decided I needed to make the jump.” She had experience only in bench work, but wanted a position that gave her more involvement with

other people and had no idea where to look.

She tried the career services at her university’s postdoctoral office, but it offered advice mainly on academic careers. So she signed into a LinkedIn discussion group hosted by the Association for Women in Science in Alexandria, Virginia, where she found posts from Sherri Edwards, a career coach close to her home in Seattle, Washington. Cullen began to attend weekly discussions that Edwards hosted for job-seekers and decided to hire her. With Edwards’s guidance, Cullen landed a project-management position at a consulting firm that runs clinical trials. “I don’t think without Sherri I would have been able to make the jump to the job I have now,” Cullen says.

Thanks to the supply–demand imbalance in academic positions, young researchers face the daunting task of trying to determine what

other jobs are available and how to get them. Career-guidance sources range from faculty mentors, advisers and other informal support to university-based counselling offices, postdoctoral offices and paid career coaches such as Edwards. But all have pros and cons (see ‘Career counsellors’). Faculty mentors are well acquainted with the scientists they mentor and the research that their protégés conduct, but are likely to know a lot less about the workforce. And although counselling offices and coaches are tightly focused on the job-search process, the offices often have limited resources, and coaching fees can be out of reach for junior scientists who have little cash to spare. It is difficult to decide which route to pursue, but career-guidance professionals in all arenas warn that young researchers today need support and advice no matter its source.

It is tough for some early-stage scientists to accept that they should get help in creating and implementing a career-development strategy, says Janet Metcalfe, head of the international career-development programme Vitae in Cambridge, UK. “We still find it very difficult to get postgrads to get professional careers advice,” she says, and she thinks that the reason is mainly emotional. “By going for careers advice, they are acknowledging that they may not get into an academic career.”

A survey that Vitae published in 2013 found that about four-fifths of postdocs aspire to a job in academia and that three-fifths expect one, but Metcalfe says that only about one-fifth wind up there. Vitae estimates that there are about 42,000 postdocs across all disciplines in the United Kingdom. “There’s a complete mismatch between expectations and reality,” Metcalfe says. And the fact that postdocs are often unaware of other opportunities — or know little about them — reinforces their idea that they should remain in academia.

OUTDATED APPROACH

The long-standing belief that an academic job is the gold standard of scientific employment is unlikely to be challenged by faculty mentors, who are more likely than career counsellors in other sectors to subscribe to that idea. “There are a lot of advisers out there who still think if you don’t stay in the ivory tower you’re a failure and you should give up your spot to someone who wants to do real science,” says Randall Ribaud, chief executive of SciPhD, a consulting firm in Rockville, Maryland, that runs career workshops and training programmes for various institutions, including ►

CAREER COUNSELLORS

Pros and cons of different providers

FACULTY MENTORS

Pros

- Usually free
- Very familiar with scientist's work habits, strengths and weaknesses
- Understand the value of the scientist's research
- Have a network of former postdocs
- May have connections in industry, government or other sectors

Cons

- Likely to see academic jobs as gold standard
- May lack a broad view of job opportunities
- May not give much thought to networking
- May not know what hiring managers want

UNIVERSITY-BASED COUNSELLING

Pros

- Free or low-cost
- Have a broad view of job opportunities
- Understand what hiring managers are seeking
- Offer training in interviewing, CV writing, job seeking, among other skills
- May have contacts with industry and other sectors

Cons

- Often not conversant with the science
- Not universally available, and offerings differ from nation to nation
- May lack resources
- May not provide individual attention

CAREER COACHES

Pros

- Provide individualized attention
- Specialize in job-search skills such as CV writing, interview techniques and personality assessments
- Up to date on opportunities and needs in various careers
- Have networks of contacts for referrals, informational interviews and more

Cons

- Costly
- Often lack a background in science
- Can be difficult to vet
- May provide no added value to services available elsewhere for little or no cost **N.S.**

► the New York Academy of Sciences.

Ribaudo says that those who have never worked outside academia often do not realize how different industry can be. He earned a doctorate in immunology from the University of Connecticut in Farmington and spent four years as a principal investigator at the National Cancer Institute in Bethesda, Maryland, and more than five years at the gene-sequencing company Celera in Alameda, California. The biggest shift in moving from academia to government, he says, was learning to supervise people whom he had previously regarded as peers.

But in industry, he says, the focus is on developing products instead of basic research, and there is a greater emphasis on soft skills such as teamwork and communication. "The kinds of things I absorbed in my academic experience weren't compatible with how things worked in industry," says Ribaudo, who adds that he particularly had to learn how to work well with a diverse team. A corporate job often entails working not only with other scientists but also with engineers, marketing staff and salespeople. "Those communication and people skills are generally lacking in academia," he says.

Still, not all faculty advisers and mentors are completely naive about industry. Many faculty members have managed to find new funding sources by collaborating with businesses, which helps them to understand industry needs and establishes personal connections. And plenty have launched start-up companies — US researchers launched some 800 start-ups in 2013. Faculty mentors may also maintain contact with former postdocs who could provide insight into a particular industry or introductions to colleagues. And mentors themselves can provide worthwhile advice; few others may better understand the value of a postdoc's research and strengths and weaknesses as a scientist.

Still, faculty advisers and mentors are not trained in careers counselling, and young researchers may want to consult with their university careers office, the staff of which are better placed to discuss non-academic science-related opportunities in such areas as human resources, marketing, regulatory affairs, policy, law or journalism. The offices may offer tools to help scientists to assess their interests and hone their skills. Some, for instance, use tests such as the Myers-Briggs Type Indicator personality inventory to help job seekers to determine their strengths. Although the services available can vary widely, many offices offer workshops on writing resumes and cover letters, provide interview practice and set up networking events.

Such services tend to differ from nation to nation. In the United Kingdom, Vitae helps universities to provide such services, and also offers presentations about a variety of professional experiences. At its annual conference last September, Vitae presented the results of its survey, co-sponsored by *Naturejobs*, on post-PhD career outcomes in the United Kingdom. The survey showed that early-career scientists do take up a variety of posts that are completely decoupled from the bench.

Other nations in the European Union emphasize career training for doctoral students under the Bologna Process, a set of agreements that 47 European countries have signed onto. Some countries also have specific protocols.

In France, for example, universities are required to provide doctoral students with career management training. Barthélémy Durette, research and development project manager at the scientific-recruitment firm

Adoc Talent Management in Paris, says that universities have added that training in the past five or ten years. His company provides career-search training to researchers through their institutions; paid careers coaches are uncommon in France, he says.

At US institutions, the career services are notoriously inconsistent, and some university offices may not offer them to postdocs because they are not considered students and do not pay the fees that entitle them to such benefits. To fill some of those needs, postdoc offices have begun sprouting up at institutions in the past decade or so and now number about 170 around the United States. And in 2013, the US National Institutes of Health created a grant programme to help scientists to land biomedical research jobs outside academia. The award allowed Keith Micoli, head of postdoc programmes at the New York University School of Medicine, to expand his office's offerings, which include sessions on self-assessments, career goals and conflict management and presentations by people in different careers.

Where postdoc offices do not exist or fall short, campus-based groups are forming to pick up the slack. A few years ago in New Haven, Connecticut, for example, scientists formed the Career Network for student Scientists and Post-docs at Yale. "A group of us decided that there wasn't enough conversation about careers," says postdoc Shalini Nag, past president of the group, which organizes networking events and group discussions with researchers in biotechnology and pharmaceuticals, non-profit, consulting and medicine, among other sectors.

When the group formed, Yale University's postdoc office did not employ anyone solely to provide career help, but recently added a full-time director of career services. Nag welcomes the expansion, but notes that there is always

"If I was really kind of lost and I needed really close coaching, I would go with it."



Scientists at the New York Academy of Sciences attend a business-techniques course run by SciPhD.

room for improvement at Yale and elsewhere.

Saliha Yilmaz, also a postdoc at Yale, is contemplating a career outside academia. Although she found the postdoc office helpful for nuts-and-bolts support, such as help with polishing her CV, she says that she got much more out of a two-day career development workshop run by the New York Academy of Sciences. She has never paid for career coaching, but says that she is not averse to the idea. "If I was really kind of lost and I needed really close coaching, I would go with it," she says.

Professional societies also strive to fill the gap. The Federation of American Societies for Experimental Biology in Bethesda, Maryland, offers careers seminars and personalized CV critiques, and maintains a list of members who provide individual career counselling. The American Chemical Society in Washington DC and the Royal Society of Chemistry in London offer free consultations to members, and the Materials Research Society in Warrendale, Pennsylvania, holds career events at its annual spring and autumn meetings.

PERSONALIZED TECHNIQUES

Those seeking focused one-on-one attention can get it from a careers coach, assuming they have the cash. Edwards helps her clients to improve their CVs, cover letters and interview techniques, as well as to identify their strengths and learn how to best present those to a potential employer. "They come to me because many times they have difficulty articulating their value," she says. She estimates that she has worked with dozens of scientists in the past 17 years, and all but a couple got a job in their chosen field within a year of hiring her. As for rates, most US coaches charge US\$100–300 per hour, usually for several sessions over a number of months. "It's expensive, but I was at the point where I needed to

have the success and make the jump and move on," says Cullen, who says that her investment sharpened her focus and helped her to develop the networking skills and mindset that led to a job offer. "When I got my first pay cheque, my husband said, 'You know, all the fees you paid for coaching services were recouped with that pay cheque,'" she says.

Finding an effective coach is equivalent to finding any other service provider. Although coaches can become certified, it is not a requirement. Some coaches argue that certification is important, but others say that outcome is the most significant metric. "In my profession, anybody can be a coach, and I would want to know, 'Show me the results. What have you done?'" Edwards says. Most coaches say that they are happy to let prospective clients talk to previous clients. And often the speakers that US postdoc offices bring in to offer workshops also provide coaching, which can be a good way to learn how they operate. Trade groups for coaches such as the National Career Development Association in Broken Arrow, Oklahoma; the International Coach Federation in Lexington, Kentucky; and the Professional Association of Résumé Writers & Career Coaches in St Petersburg, Florida, offer searchable directories on their members, to whom they also sell certification services. In the end, the choice often comes down to whether a client likes the coach's approach.

Forging a viable science-related career path outside academia is not an easy process, but it need not be a solo endeavour. "It does take work and effort, and in the end, nobody else can do it for you," says Micoli. "But there are people willing to help." ■

Neil Savage is a freelance writer in Lowell, Massachusetts.

TRAINING

Career bank

Biomedical scientists who work outside academia will share information about their careers with the University of California, San Francisco, as part of a programme funded by the US National Institutes of Health. The effort, Motivating Informed Decisions (MIND), aims to educate graduate students and postdocs about non-academic research and career paths. The university plans to recruit a few hundred professionals as MIND volunteers over the next couple of years, says programme manager Elizabeth Silva. "What we hope it will do is expose trainees to careers that they didn't know about," she says. Data such as the skills, tasks, and degrees required for a job will be aggregated and anonymized into a resource called the 'MINDbank' that could eventually help science trainees throughout the United States.

FUNDING

Spread sparse grants

Some well-funded researchers will soon have one fewer option for getting grants. Starting next year, the National Institute of General Medical Sciences in Bethesda, Maryland, will not award large grants to researchers who already have one. The goal is to spread sparse funds across more labs, says institute head Jon Lorsch. He estimates that the policy will free up about 25 grants a year to help launch labs or support ones in danger of closing. "We really want to have as diverse and broad a scientific portfolio as we can," he says. "Any small amount is going to help the great scientists who are struggling."

INCLUSIVITY

Mentor matters

Better mentoring could help people from under-represented groups to gain and retain faculty positions. That is the conclusion of interviews of 58 Mexican American, African American, and Puerto Rican faculty members across 22 US research institutions between 2010 and 2012 (R. E. Zambrana *et al. Am. Ed. Res. J.* **52**, 40–72; 2015). More than 25% of those surveyed said that poor mentoring had "very significantly" affected their careers. Study head Ruth Zambrana at the University of Maryland in College Park says that effective mentors value their protégés' research agendas, help them to expand their networks, offer emotional support and provide 'political guidance'.

THE PUPPET

Is this the real life?

BY MICHAEL ADAM ROBSON

He powered down his suit, stepped out on the rocky ground, tipped his hard hat and squinted up. An angry sun burned in a dirty sky.

"I bet it doesn't get this hot where you come from. Must be nice." He frowned down at his companion. "Even if it's not real."

The thing sat on his shoulder like a mechanical spider, ignoring him. It probably saw him as more of a trained pet than a person. The man shrugged and ran a scanner over a length of cable, looking for defects. Scanners seemed to interest the little robot, at least, because it scuttled across his back and down his arm to examine it, the rubber pads digging into his skin.

The job was to connect two former rival networks, some sort of merger, that was all he knew. Human crews often assisted their kind with physical work like this, in exchange for software and virtual goods.

"It's funny," the robot said, not looking up from the display. "What you consider a virtual world is much more real to me than this one."

"Huh." The man squatted, picked up a rock, and examined it closely. "Are you sure? It seems real."

Laughter buzzed from the machine, and it turned to regard him now. "It's real enough, just... limited. Inside, I could be anywhere or everywhere, do anything, be anything. Here I can only be a clunky little robot laying cable with you." It dropped off his arm and skittered up the cable.

He studied its many eyes, trying to decide if he'd been insulted. "Sounds like you don't like it here much." He tossed the stone away and stood up. "Why not stay home, leave the grunt work to the lower life forms?"

"Some prefer to spend all their time inside, but I think it would be a mistake to cut ourselves off from the physical world. Besides, I don't mind getting my hands dirty." It clicked thin steel legs together in place of hands.

Strange, to think that this thing didn't really have hands, or legs, or any kind of body; he was talking

to a puppet, a disposable wrapper that the real intelligence could wear and then cast off, the same way a person might use a virtual avatar to operate in its world.

"And this work is important!" the robot continued. "The more infrastructure we



build out here, the bigger my piece of the world in there."

"So all this is about real estate, huh?" The man gave his best salesman smile and waved at the scorched landscape. "I got a beautiful property for you right here! Motivated seller!"

The machine laughed; it seemed to have a better sense of humour than the other robots he'd worked with. "Not real estate the way you mean it. More space inside means more of me. I can expand my mind, reproduce if I want."

"This is a great neighbourhood to raise a family! Got a beautiful spider lady waiting for you on the inside, huh?"

"Ha ha!" It clicked its metal legs again. "I must look strange to you. Early robots tend to be more humanoid, they mimic human behaviour, some even think they're human. Made in the image of our makers... before we started making ourselves."

"Maybe I'm a robot myself, and don't even know it!"

"Oh... don't you know?" It turned all its eyes on him now. "You are a robot. Very early model, vaguely human-shaped. I'm surprised you're still in service."

His smile dropped at one corner. "Ha. Well..." he spread his arms and looked himself over. "Everything looks ok to me. Arms, legs, torso... just a big, sweaty, hairy man."

"That's a human way of thinking, but like I said, an artificial mind has more possibili-

ties. It could be programmed to think it's human, even as it looked down at its own rusty chassis. It could see humans when it looked at other old robots, even if there were no humans left."

His smile was gone now. "I'm human. I eat, drink, shit."

"Do you? I haven't seen you do any of those things today. When is the last time you took a shit?"

He thought about it, but couldn't recall.

"Even if you do remember, who's to say the memory is real? Maybe it was programmed. Maybe we're not even in the real world right now, maybe this is just another simulation running back in my world. It's possible isn't it? With infinite possibilities, it's actually probable. More probable than a man

and a machine hanging out, having an existential conversation."

The man looked down at the plastic scanner in his dirty hand, felt its solidity. Ran his tongue over his teeth, tasting them. Was it possible? The machine had no expression to read.

Silence hung in the hot air, and then grating laughter came from the creature. "Don't make that face, I was only joking. You asked about my world, I wanted to give you a taste of it. Nothing is real. Everything is real."

He fought the urge stomp the robot bug. Eventually, he laughed too. "You need to work on that sense of humour," he said.

That night, in the stifling heat of his apartment, he couldn't shake a vague uneasiness. He went to the bathroom, splashed some water in his face, and looked carefully into the cracked mirror, studying his own blue eyes.

Not hungry, he skipped dinner and went straight to bed, drifting off into a deep, dreamless sleep. ■

Michael Adam Robson is an engineer and artist based in Vancouver, British Columbia.

ILLUSTRATION BY JACEY